OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: September 23, 2005, 16:33:30 ; Search time 3423 Seconds
                 (without alignments)
                 15968.534 Million cell updates/sec

Title: US-10-602-166-2
Perfect Score: 1436
Sequence: 1 cattcagaatcatctccagc.........tatcaggtggttcaacctgg 1436
Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0
Searched: 34239544 seqs, 1903134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   EST:*
                 1: gb_est1:*
                 2: gb_est2:*
                 3: gb_est3:*
                 4: gb_htc:*
                 5: gb_est4:*
                 6: gb_est5:*
                 7: gb_est6:*
                 8: gb_gss1:*
                 9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	307.2	21.4	376	9	CL921501	CL921501 OA_ABa0002
c	2	130.8	9.1	612	9	CL835542	CL835542 OR_CBa005
c	3	126	8.8	784	9	CL825571	CL825571 OR_CBa008
c	4	126	8.8	829	9	CL855333	CL855333 OR_CBa008
	5	124	8.6	840	9	CL728092	CL728092 OR_BBa005
	6	123.6	8.6	781	9	CL783664	CL783664 OR_BBa010
	7	119.8	8.3	486	9	CL811645	CL811645 OR_BBa002
	8	107.8	7.5	665	9	CL708436	CL708436 OR_BBa002
	9	107.8	7.5	820	9	CL854157	CL854157 OR_CBa008
	10	107.8	7.5	824	9	CL782578	CL782578 OR_CBa009
c	11	105.2	7.3	784	9	CL746481	CL746481 OR_BBa008
c	12	104.6	7.3	707	9	CL807668	CL807668 OR_BBa002
	13	104.6	7.3	757	9	CL843536	CL843536 OR_BBa007
c	14	104	7.2	590	8	AQ328957	AQ328957 nbxb0044I
	15	100.6	7.0	431	9	CG206879	CG206879 TOS0501T
	16	99.6	6.9	651	9	CW508204	CW508204 OP_Ba000
	17	99	6.9	387	8	AQ912524	AQ912524 nbeb0037I
	18	99	6.9	845	8	AZ045366	AZ045366 nbeb0080N
c	19	98.8	6.9	771	9	CL753477	CL753477 OR_BBa012
c	20	98.4	6.9	754	9	CL858870	CL858870 OR_CBa008
c	21	97	6.8	723	9	CL817465	CL817465 OR_CBa003
c	22	94.4	6.6	491	9	CL739488	CL739488 OR_BBa007
c	23	93.4	6.5	365	6	C72467	C72467 C72467 Rice
	24	92.2	6.4	561	9	CL752715	CL752715 OR_BBa011
	25	91.4	6.4	280	6	C71788	C71788 C71788 Rice
	26	91.2	6.4	642	8	AQ578656	AQ578656 nbxb0093C
	27	90.2	6.3	624	9	CL790230	CL790230 OR_BBa010
	28	90.2	6.3	640	9	CL749018	CL749018 OR_BBa011
	29	90.2	6.3	774	9	CL612072	CL612072 OR_BBa000
	30	90	6.3	562	9	CL629198	CL629198 OR_BBa002
	31	88.4	6.2	754	8	AQ840531	AQ840531 nbxb0063P
	32	86.2	6.0	699	9	CL858276	CL858276 OR_CBa008
	33	86.2	6.0	773	9	CL736927	CL736927 OR_BBa007
	34	86	6.0	742	9	CL627803	CL627803 OR_BBa002
	35	84.6	5.9	738	9	CL753761	CL753761 OR_BBa012
	36	83.2	5.8	789	9	CL768364	CL768364 OR_BBa005
	37	82.8	5.8	621	9	CW538474	CW538474 OP_Ba005
	38	82.6	5.8	617	9	CL853543	CL853543 OR_CBa008
	39	81.2	5.7	745	9	CL838261	CL838261 OR_CBa006
	40	81	5.6	336	9	AG212556	AG212556 Oryza sat
	41	81	5.6	576	9	AG212556	AG212556 Oryza sat
c	42	80.6	5.6	756	9	CL859905	CL859905 OR_CBa009
c	43	80.6	5.6	677	9	CL772037	CL772037 OR_BBa008
c	44	80	5.6	490	9	AB156569	AB156569 Oryza sat
c	45	80	5.6	685	9	AG021430	AG021430 Oryza sat

ALIGNMENTS

RESULT 1
CL921501/c
LOCUS   CL921501   376 bp   DNA   linear   GSS 14-SEP-2004
DEFINITION OA_ABa0026D24.r OA_ABa Oryza australiensis genomic clone
            OA_ABa0026D24 3', genomic survey sequence.
ACCESSION   CL921501
VERSION   CL921501.1 GI:52037176
KEYWORDS   GSS.
SOURCE   Oryza australiensis
  ORGANISM Oryza australiensis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1 (bases 1 to 376)
  AUTHORS   Kim,H., Yu,Y., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R.,
            Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
  TITLE   OMAP Project
  JOURNAL   Unpublished (2004)
COMMENT   Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: http://genome.arizona.edu
            PCR Primers
            FORWARD: TAA TAC GAC TCA CTA TAG GG
            BACKWARD: CAC TCA TTA GGC ACC CCA
            Plate: 0026   row: D column: 24
            Seq primer: CAC TCA TTA GGC ACC CCA
            Class: BAC ends.
FEATURES   Location/Qualifiers
     source   1..376
                     /organism="Oryza australiensis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4532"
                     /clone="OA_ABa0026D24"
                     /tissue_type="young leaves"
                     /lab_host="DH10B T1 phage resistant"
                     /clone_lib="OA_ABa"
                     /note="Vector: pACiBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN

Query Match   21.4%; Score 307.2; DB 9; Length 376;
Best Local Similarity   94.1%; Pred. No. 2e-66;
Matches 354; Conservative   0; Mismatches   13; Indels   9; Gaps   3;

RESULT 2

```
Qy   510 AATGTGCCATGATCCTGCACATCGCCAAGATTATATTCCTTCACATTTTTCTTCC 569
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   376 AATGTGCCATGATCCTGCACATCGCCAAGATTATATTCCTTCACATTTTTCTTCC 317

Qy   570 TATCGCTCCTAGTC-------GTCTGTTTGGGGAGCTTAAAAATTATTCCTCACATT 622
         |||||||||||||       |||||||||||||||||||||||||||||||||
Db   316 TATCGCTCCTAGTCTTCTATGGTCTGTTTGGGGAGCTTAAAAATTATGAAAAGCAGCTGCTG 257

Qy   623 AGAAGCTAGCTGTGTGAGAATCTGAAGAATTGAGTTCTAGGTTCATTCTCCAGATTCTAC 682
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   256 AGAAGTTAGCTGTGTGAGAATCTGAAGAATTAGGTTCTA-GTTCATTCTCCAGATTCTAC 198

Qy   683 AATTACAGATTCTTATATATTTAGGTAAAAAGTCGGACTGTTT-GGGAGCTTCTGTCAGCC 741
         |||||||||||||  ||||||||||||||||||||||||   ||||||||||||||
Db   197 AATTACAGATTCTTAGAATTAGGTAAAAAAGCTGGAACTGTTTGGGGAGCTTCTGTCAGCC 138

Qy   742 GGAGATTCTGTGAGAAGCGTGCGAGCTTCCCAAACAGACGACCCCTAGTTGTTAC 801
         ||||||||||||||||||||  ||||||||||||||||||||||||||
Db   137 GGAGATTCTGGGAGAAGCTGCCAACTGGAGAAGCTTCCCAAAACAGACCCCTAGTTGTTAC 78

Qy   802 TCTAGCTGATCGATTCACTCTATTATTATATACACCTTGGTCTCTTAGCTTATCAAACGTTAG 861
         |||||||||||||||||||  ||||||||||||||||||||||||||
Db   77  TCCAGCTGATCGATTCACTCTATTTTATATTGCACTTTGCTCTTAGCTTATCAAACGTTAG 18

Qy   862 CCAAGACTTGAATTT 877
         |||||||||||||||
Db   17  CCAAGACTTGAAATTT 2
```

RESULT 3

```
Qy   581 GTCGTCTGTTTGGGGAGCTTAAAAATATTATGAAAAGCAGCTGCTGAGAAGCTAGCTGGTGAGA 640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   433 GGCCTGTTTGGGGGGAGATTAAGATTCCGAGAAGCTGCTGTTGAGAAGCTAGCTGGTGAGA 374

Qy   641 ATCTGAAGAATTTG-----AGTTCTACGTTCATTCTCCAGATTCTACAACAATTACACGATTCTT 696
         ||||||||||||    |||||||||||||||||||||||||||||||||||
Db   373 ATCTAGAGAAGCTGGGAAAGCTTCTAGTTCATTTCGCAGATTCTACAACAATTACACGATTCTC 314

Qy   697 ATAATTTAGGTAAAAAGTCGGACTGTTT--GGGGAGCTTCTGTCACCCGGAGATTCTGTGA 754
         ||||||||||||||||||||||||||   ||||||||||||||||||||||
Db   313 AGAATCTGGTAAAAAAGTTGGACTGTTTGGGGGGAGCTTCTCGGGAACTGGAGATTCTAGGA 254

Qy   755 GAAGCTGCAGCTGCTGAGAAGCTTCCCAAAACAGACGACCCCTAGTTGTACTCTAGCTGATCGA 814
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   253 GAAGCTGCAGCTGCTGAGAAGCTCCCCAAAACAGGCGCCCCTTAGGGTTCGGCCGACAGTGAGC 194

Qy   815 TTCACTCTATTTTA 828
         ||||||||||||||
Db   193 TTTTCCAGCTTATA 180
```

Query Match   8.8%; Score 126; DB 9; Length 784;
Best Local Similarity   80.1%; Pred. No. 1.4e-20;
Matches 173; Conservative   0; Mismatches 40; Indels   3; Gaps   2;

```
Qy   581 GTCGTCTCTGTTTGGGGAGCTTAAAATTATGAAAAGACAGCTGCTGAGAAGCTAGCTGGTGAGA 640
             ||||| |||| | || ||||||||||||||||||||||||||| ||| ||| |||||||||
Db   391 GACCTGTTTGGGGGAGGTTAAGATTCTGAAAAAGCAGCCCCTTAAGAAGCTAGCTCGTGAGA 450

Qy   641 ATCTGAAGAAATTGAGTTCTACGTTCATTCTCCAGATTCTACAATTACAGATTCTTATAA 700
             | || |||||  |||| ||  |||||||||||||||||||| || ||| ||| |||
Db   451 AACCCAGC--TTCTGGCTTCTAGTTCATTTTCTAGATTCTTACAACTACAGATTCTTAGAA 508

Qy   701 TTTAGGTAAAAAGCTGGGACTGTTTGGG--AGCTTCTGTCGTCGCCGGAGAATTCTGTGAGAAGC 759
             ||||||||||| | || ||||||||||   ||||| | |||| |||| |||| ||| || ||
Db   509 TCTAGGTAAAAATGCTGGGACTGTTTGGGCAACTGGAGAATTCTTAGAGGAAGC 568

Qy   760 TGCAGCTGCTAGAAGCTTCCCCAAACAGACCCCTAG 795
             ||||| |||||||||| |||||| ||||| ||||||
Db   569 TGCAGTTGCTAGAAGCTCCCCCAAAATAGGCCCTAG 604
```

RESULT 4

LOCUS   CL855333   829 bp   DNA   linear   GSS 16-AUG-2004
DEFINITION OR_CBa0008SC07.f OR_CBa Oryza rufipogon genomic clone OR_CBa0008SC07
            5'_ genomic survey sequence.
ACCESSION   CL855333
VERSION   CL855333.1 GI:51263573
KEYWORDS   GSS.
SOURCE   Oryza rufipogon
  ORGANISM Oryza rufipogon
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1 (bases 1 to 829)
  AUTHORS   Kim,H., Yu,Y., Wissotski,M., Yost,D., Stum,D., Rao,K., Luo,M.,
            Jetty,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and
            Wing,R.
  TITLE   OMAP project
  JOURNAL   Unpublished (2004)
COMMENT   Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: http://genome.arizona.edu
            PCR Primers
            FORWARD: TAA TAC GAC TCA CTA TAG GG
            BACKWARD: CAC TCA TTA GGC ACC CCA
            Plate: 0085 row: C column: 07
            Seq primer: TAA TAC GAC TCA CTA TAG GG
            Class: BAC ends.
FEATURES   Location/Qualifiers
     source   1..829
                     /organism="Oryza rufipogon"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4529"
                     /clone="OR_CBa0008SC07"
                     /tissue_type="young leaves"
                     /dev_stage="2 week old seedlings"
                     /lab_host="DH10B T1 phage resistant"
                     /clone_lib="OR_CBa"
                     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII;
                     drk treated 36 hrs before harvest"
ORIGIN

Query Match   8.8%; Score 126; DB 9; Length 829;
Best Local Similarity   80.1%; Pred. No. 1.4e-20;
Matches 173; Conservative   0; Mismatches 40; Indels   3; Gaps   2;

```
Qy   581 GTCGTCTCTGTTTGGGGAGCTTAAAATTATGAAAAGACAGCTGCTGAGAAGCTAGCTGGTGAGA 640
```

Db   393 GACCTGTTTGGGGGAGGTTAAGATTCTGAAAAAGCAGCCCCTTAAGAAGCTAGCTCGTGAGA 452
             ||||| |||| | || ||||||||||||||||||||||||||| ||| ||| |||||||||
Qy   641 ATCTGAAGAAATTGAGTTCTACGTTCATTCTCCAGATTCTACAATTACAGATTCTTATAA 700

Db   453 AACCCAGC--TTCTGGCTTCTAGTTCATTTTCTAGATTCTTACAACTACAGATTCTTAGAA 510
Qy   701 TTTAGGTAAAAAGCTGGGACTGTTTGGG--AGCTTCTGTCGTCGCCGGAGAATTCTGTGAGAAGC 759

Db   511 TCTAGGTAAAAATGCTGGGACTGTTTGGGCAACTGGAGAATTCTTAGAGGAAGC 570
Qy   760 TGCAGCTGCTAGAAGCTTCCCCAAACAGACCCCTAG 795

Db   571 TGCAGTTGCTAGAAGCTCCCCCAAAATAGGCCCTAG 606

RESULT 5

LOCUS   CL728092   840 bp   DNA   linear   GSS 27-JUL-2004
DEFINITION OR_BBa0059A19.f OR_BBa Oryza rufipogon genomic clone OR_BBa0059A19
            5'_ genomic survey sequence.
ACCESSION   CL728092
VERSION   CL728092.1 GI:50662448
KEYWORDS   GSS.
SOURCE   Oryza rufipogon
  ORGANISM Oryza rufipogon
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1 (bases 1 to 840)
  AUTHORS   Kim,H., Yu,Y., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R.,
            Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
  TITLE   OMAP project
  JOURNAL   Unpublished (2004)
COMMENT   Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: http://genome.arizona.edu
            PCR Primers
            FORWARD: TAA TAC GAC TCA CTA TAG GG
            BACKWARD: CAC TCA TTA GGC ACC CCA
            Insert Length: 161   Std Error: 0.00
            Plate: 0059 row: A column: 19
            Seq primer: TAA TAC GAC TCA CTA TAG GG
            Class: BAC ends.
FEATURES   Location/Qualifiers
     source   1..840
                     /organism="Oryza rufipogon"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4529"
                     /clone="OR_BBa0059A19"
                     /tissue_type="young leaves"
                     /lab_host="DH10B-T1 phage resistant"
                     /clone_lib="OR_BBa"
                     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN

Query Match   8.6%; Score 124; DB 9; Length 840;
Best Local Similarity   84.1%; Pred. No. 4.4e-20;
Matches 175; Conservative   0; Mismatches 30; Indels   3; Gaps   3;

Qy   589 TTTGGGGAGCTTAAAATTATGAAAAGACAGCTGCTGAGAAGCTAGCTGGTGAGAAGCTGGTGAGA 648
             ||||||||| | | ||||| || |||||||||||||||||||||||||||||||||
Db   436 TTGGGGGAGGTTAAGATTCTGAAAAAGCAGCTGCTAGGAAGCTAGCTGGTGAGAAAC-CCAG 494

Qy   649 AATTTGAGTTCTACGTTCATTCTCCAGATTCTACAATTACAGATTCTTATAATTAGGTA 708
             || ||| |||||| |||||||||||||||||||| || ||||||||| || |||||||
Db   495 CTTCTGACTTCTA-GTTTATTTTCTAGATTCTACAACTACAGATTCTTAGAATATAGGTA 553

Qy   709 AAAAGCTGGGACTGTTTGGG--AGCTTCTGTCGTCGCCGGAGATTCTGTGAGAAGCTGCAGCTG 767
             |||||||||||||||||||| |||| | |||| |||| ||||| || || |||
```

RESULT 6

```
Db    CL783664
Qy    768 CTAGAAGCTTCCCAAACAGACCCCTAG 795
          ||||||||||||||||||||||||||||
Db    614 CTAGAAGCTCCCCAAATAGGCCCTTAG 641
```

```
LOCUS       CL783664              781 bp    DNA    linear   GSS 30-JUL-2004
DEFINITION  OR_BBa0100G05.f OR_BBa Oryza rufipogon genomic clone OR_BBa0100G05
            5',genomic survey sequence.
ACCESSION   CL783664
VERSION     CL783664.1  GI:50863315
KEYWORDS    GSS.
SOURCE      Oryza rufipogon
  ORGANISM  Oryza rufipogon
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 781)
  AUTHORS   Kim,H., Yu,Y., Stum,D., Rao,K., Luo,M., Jetty,R.,
            Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
  TITLE     OMAP Project
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: http://genome.arizona.edu
            PCR Primers
            FORWARD: TAA TAC GAC TCA CTA TAG GG
            BACKWARD: CAC TCA TTA GGC ACC CCA
            Insert Length: 161     Std Error: 0.00
            Plate: 0100   row: G     column: 05
            Seq primer: TAA TAC GAC TCA CTA TAG GG
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..781
                     /organism="Oryza rufipogon"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4529"
                     /clone="OR_BBa0100G05"
                     /tissue_type="young leaves"
                     /lab_host="DH10B-T1 phage resistant"
                     /clone_lib="OR_BBa"
                     /note="vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN
```

Query Match   8.6%; Score 123.6; DB 9; Length 781;
Best Local Similarity 84.5%; Pred. No. 5.5e-20;
Matches 174; Conservative 0; Mismatches 29; Indels 3; Gaps 3;

```
Qy  589 TTTGGGAGCTTAAAATTATGAAAAGCAGCTAGCTGGTGGTCGAGAGATCTGGAG 648
        ||||| ||||||||| || ||||| ||||||||||| |||  ||| | ||||||
Db  436 TTGGGGAGGTTAAGATTCAGAAAAGCAGCTGCTRAGAGAAGCTAGCTGGTCGAGAAAC-CCAG 494

Qy  649 AATTTGAGTTCTACGTTCATTCTCCAGAATTCTACATTACAGAATTCTTATATAATTAGGTA 708
        |||||||| ||| |||| || || ||||||||||| || |||| ||| ||| ||| ||||
Db  495 CTTCTGACTTCTA-GTTTATTTTCTAGATTCTACAACTACAGAATTCTTAGAATATAGGTA 553

Qy  709 AAAAGCTGGACTGTTTGGG-AGCTTCGTGTCAGCCGGAGATTCTGTGGAGAAGCTGCAGCTG 767
        |||||||||| ||||| || ||| || | |||| |||||| ||||||||||||||||||| ||
Db  554 AAAAGCTGGACTGTTTGGGAGCTTCTAGCAACTGAAGATTCTAGGAGAAGCTCGCAGTTG 613

Qy  768 CTAGAAGCTTCCCAAACAGACCCCT 793
        ||||||||| |||||||| ||||||
Db  614 CTAGAAGCTCCCCAAATAGGCCCCT 639
```

RESULT 7

```
LOCUS       CL811645              486 bp    DNA    linear   GSS 09-AUG-2004
DEFINITION  OR_CBa0027D24.r OR_CBa Oryza rufipogon genomic clone OR_CBA0027D24
            3',genomic survey sequence.
ACCESSION   CL811645
VERSION     CL811645.1  GI:51051839
KEYWORDS    GSS.
SOURCE      Oryza rufipogon
  ORGANISM  Oryza rufipogon
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 486)
  AUTHORS   Kim,H. Yu,Y. Wissotski,M., Yost,D., Stum,D. Rao,K., Luo,M.,
            Jetty,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and
            Wing,R.
  TITLE     OMAP project
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: http://genome.arizona.edu
            PCR Primers
            FORWARD: TAA TAC GAC TCA CTA TAG GG
            BACKWARD: CAC TCA TTA GGC ACC CCA
            Plate: 0027   row: D     column: 24
            Seq primer: CAC TCA TTA GGC ACC CCA
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..486
                     /organism="Oryza rufipogon"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4529"
                     /clone="OR_CBa0027D24"
                     /dev_stage="2 week old seedlings"
                     /tissue_type="young leaves"
                     /lab_host="DH10B T1 phage resistant"
                     /clone_lib="OR_CBa"
                     /note="vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII;
                     drk treated 36 hrs before harvest"
ORIGIN
```

Query Match   8.3%; Score 119.8; DB 9; Length 486;
Best Local Similarity 72.1%; Pred. No. 4.5e-19;
Matches 191; Conservative 0; Mismatches 57; Indels 17; Gaps 2;

```
Qy  581 GTCGTCTGTTTGGGAGCTTAAAATTATGAAAAGCAGCTCGCTGAGAAGCTTAGCTGGTCGGTGGTGAGA 640
        ||||||||| |||||||| |||| ||| || |||||| | |||||||||| |||| |||| ||| |||
Db    8 GGGCCTGTTTGGGGGGAGCTTAAGATTTGAGAAGCAGCTGCTGAGAAGCTAGCCGGTGAGA 67

Qy  641 ATCTGAAGAATTGA------------------GTTCTACGTTCATTCTCCAGATTCTACAAT 685
        || | ||| | |||                  ||||||| ||||||| || |||||| |||
Db   68 ATCTGGAGAAGCTGGAAAACCCAGCTTCTGGCTTCTAGTTCATTTTCAGATTCTTACAAC 127

Qy  686 TACAGAATTCTTATATAATTTAGGTAAAAAAGCTGGACTGTTT--GGGAGCTTCGTGTCAGCCGCG 743
        ||||||||||| ||| ||||||| |||||||||||||||||||  ||||| ||  | |||| || |
Db  128 TACAGATTCTTAATTTAGGTAAAAAAGCTGGACTGTTTGGGAGCTTCTGTGTCAGCGCCGG 187

Qy  744 AGATTCTGTGGAGAAGCTGCAGCTGCTAGAAGCTTCCCAAAACAGACCCCTTAGTTGTACTC 803
        |||||| |||||||| | ||| |||||||||||| |||||||||||||||||||| | |||||
Db  188 AGATTCTACTAGTAGAAGCTACAGCTGCTAGAGAGCTCCCCAAAACAGACGCCCATATGACAACCT 247

Qy  804 TAGCTGATCGGATTCACTCTATTTTA 828
        |||||||| |||||| || ||| |||
Db  248 ACTCTCCTCTTTGGGCTGCAGTTAA 272
```

RESULT 8

```
LOCUS       CL708436              665 bp    DNA    linear   GSS 26-JUL-2004
```

RESULT 9

```
DEFINITION   OR_BBa0029A21.r OR_BBa Oryza rufipogon genomic clone OR_BBa0029A21
             3', genomic survey sequence.
ACCESSION    CL708436
VERSION      CL708436.1  GI:50595474
KEYWORDS     GSS.
SOURCE       Oryza rufipogon
  ORGANISM   Oryza rufipogon
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
             Ehrhartoideae; Oryzeae; Oryza.
REFERENCE    1  (bases 1 to 665)
  AUTHORS    Kim,H., Yu,Y., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R.,
             Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
  TITLE      OMAP Project
  JOURNAL    Unpublished (2004)
COMMENT      Contact: Rod A. Wing
             Arizona Genomics Institute
             University of Arizona
             Forbes Building Room 303, Tucson, AZ 85721-0036, USA
             Tel: 520 626 9595
             Fax: 520 621 1259
             Email: http://genome.arizona.edu
             PCR Primers
             FORWARD: TAA TAC GAC TCA CTA TAG GG
             BACKWARD: CAC TCA TTA GGC ACC CCA
             Insert Length: 161   Std Error: 0.00
             Plate: 0029  row: A  column: 21
             Seq primer: CAC TCA TTA GGC ACC CCA
             Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..665
                     /organism="Oryza rufipogon"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4529"
                     /clone="OR_BBa0029A21"
                     /tissue_type="young leaves"
                     /lab_host="DH10B-T1 phage resistant"
                     /clone_lib="OR_BBa"
                     /note="vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN

Query Match     7.5%;  Score 107.8;  DB 9;  Length 665;
Best Local Similarity   73.0%;  Pred. No. 5.2e-16;
Matches 173; Conservative  0; Mismatches 47; Indels 17; Gaps  2;

Qy   580 AGTCGTCTGTTTGGGGAGCTTAAAATATGAAAAGCAGCTGCTGAGAAGCTAGCTGGTGAG 639

Db   355 AGCCCTGTTTAGGGGAGCTTGAGAGATTCTAAGAGAAGCTTCACCTGGTGAG 393

Qy   640 AATCTGAAGAAATTTGAGT----------- 684

Db   394 AATCTGAAAAAGCTGAGAAACCTAGCTTCTCAGTTCATTTTACAGATTCTACAA 414

Qy   685 TTACAGATTCTTATAATTTAGGTAAAAAAGCTGAGACTGTTT--GGGAGCTTCTGTCAGCCG 742

Db   415 CTACATATTTTCAGAATCTAAGTAAATAGCTGGACTGGTGAGACTGTTTCTGCCAACTG 474

Qy   743 GAGATTCTGTGAGGAGGCTGCAGCTGCTAGAGAGCTTCCCCAAACAGACCCCTAGTTGT 799

Db   475 AAGATTCTAGGAGGAAGCTCCAGCTGTTAGGAAACTCCACCAAACAGACCCTTAACAGT 535
```

RESULT 10

```
LOCUS        CL854157                 820 bp   DNA    linear   GSS 16-AUG-2004
DEFINITION   OR_CBa0083J07.f OR_CBa Oryza rufipogon genomic clone OR_CBa0083J07
             5', genomic survey sequence.
ACCESSION    CL854157
VERSION      CL854157.1  GI:51260708
KEYWORDS     GSS.
SOURCE       Oryza rufipogon
  ORGANISM   Oryza rufipogon
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
             Ehrhartoideae; Oryzeae; Oryza.
REFERENCE    1  (bases 1 to 820)
  AUTHORS    Kim,H., Yu,Y., Wissotski,M., Yost,D., Stum,D., Rao,K., Luo,M.,
             Jetty,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and
             Wing,R.
  TITLE      OMAP project
  JOURNAL    Unpublished (2004)
COMMENT      Contact: Rod A. Wing
             Arizona Genomics Institute
             University of Arizona
             Forbes Building Room 303, Tucson, AZ 85721-0036, USA
             Tel: 520 626 9595
             Fax: 520 621 1259
             Email: http://genome.arizona.edu
             PCR Primers
             FORWARD: TAA TAC GAC TCA CTA TAG GG
             BACKWARD: CAC TCA TTA GGC ACC CCA
             Plate: 0083  row: J  column: 07
             Seq primer: CAC TCA GAC TCA CTA TAG GG
             Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..820
                     /organism="Oryza rufipogon"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4529"
                     /clone="OR_CBa0083J07"
                     /tissue_type="young leaves"
                     /dev_stage="2 week old seedlings"
                     /lab_host="DH10B T1 phage resistant"
                     /clone_lib="OR_CBa"
                     /note="vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII;
                     drk treated 36 hrs before harvest"
ORIGIN

Query Match     7.5%;  Score 107.8;  DB 9;  Length 820;
Best Local Similarity   73.0%;  Pred. No. 5.4e-16;
Matches 173; Conservative  0; Mismatches 47; Indels 17; Gaps  2;

Qy   580 AGTCGTCTGTTTGGGGAGCTTAAAATATGAAAAGCAGCTGCTGAGAAGCTAGCTGGTGAG 639

Db   334 AGCCCTGTTTAGGGGAGCTTGAGAGATTCTAAGAGAAGCTTCACCTGGTGAG 393

Qy   640 AATCTGAAGAAATTTGAGT----------- 684

Db   394 AATCTGAAAAAGCTGAGAAACCTAGCTTCTCAGTTCATTTTACAGATTCTACAA 453

Qy   685 TTACAGATTCTTATAATTTAGGTAAAAAAGCTGAGACTGTTT--GGGAGCTTCTGTCAGCCG 742

Db   454 CTACATATTTTCAGAATCTAAGTAAATAGCTGGACTGGTGAGACTGTTTCTGCCAACTG 513

Qy   743 GAGATTCTGTGAGGAGGCTGCAGCTGCTAGAGAGCTTCCCCAAACAGACCCCTAGTTGT 799

Db   514 AAGATTCTAGGAGGAAGCTCCAGCTGTTAGGAAACTCCACCAAACAGACCCTTAACAGT 570
```

```
LOCUS        CL782578                 824 bp   DNA    linear   GSS 30-JUL-2004
DEFINITION   OR_BBa0098L11.f OR_BBa Oryza rufipogon genomic clone OR_BBa0098L11
             5', genomic survey sequence.
ACCESSION    CL782578
VERSION      CL782578.1  GI:50862174
KEYWORDS     GSS.
SOURCE       Oryza rufipogon
  ORGANISM   Oryza rufipogon
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
             Ehrhartoideae; Oryzeae; Oryza.
REFERENCE    1  (bases 1 to 824)
  AUTHORS    Kim,H., Yu,Y., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R.,
             Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
  TITLE      OMAP Project
```

JOURNAL   Unpublished (2004)
COMMENT   Contact: Rod A. Wing
           Arizona Genomics Institute
           University of Arizona
           Forbes Building Room 303, Tucson, AZ 85721-0036, USA
           Tel: 520 626 9595
           Fax: 520 621 1259
           Email: http://genome.arizona.edu
           PCR Primers
           FORWARD: TAA TAC GAC TCA CTA TAG GG
           BACKWARD: CAC TCA TTA GGC ACC CCA
           Insert Length: 161   Std Error: 0.00
           Plate: 0098   row: L   column: 11
           Seq primer: TAA TAC GAC TCA CTA TAG GG
           Class: BAC ends.

FEATURES   Location/Qualifiers
     source   1..824
                     /organism="Oryza rufipogon"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4529"
                     /clone="OR_BBa0098L11"
                     /tissue_type="young leaves"
                     /clone_lib="OR_BBa"
                     /lab_host="DH10B-T1 phage resistant"
                     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match
Best Local Similarity   73.0%;   Score 107.8; DB 9; Length 824;
Matches 173; Conservative 0; Mismatches 47; Indels 17; Gaps 2;

Qy   589 TTTGGGGAGCTTAAAATTATGAAAAAGCAGCTAGCTGGTGGTGAGAAGCTGAAG 648
         ||||||||||||| ||| | ||||||||||| || |||||||| | ||||
Db   517 TTCGGGGAGCTTAAAAATTCTAAGAAAACAGCTACGTGGTGAGAACATGAAA 458

Qy   649 AATTTGAGTTCTAC----------GTTCATTCTCCAGATTCTACAATTACACAGATT 693
         | ||| | | | ||||| | | ||||||||||| | || | ||||
Db   457 AAGCTGGGAAACTCAACTTCTCGGGTTCATTTCCAGATTCTACAACTATAGATT 398

Qy   694 CTTATAATTAGGTAAAAAGCTGGACTGTTT--GGGAGCTTCGTGTCAGCCGGAGATTCTG 751
         || | | | |||| |||||||||| ||||| |||||||| | | || ||||||||
Db   397 CTCAGAATCTGGGATAAAAAGCTGGACAGTTTGGGGGAGCTCCTAGCAACTCGAGATTCTA 338

Qy   752 TGAGAAGCTGCAGCGTGCTTAGAAGCTTCCCCAAACAGACCCCCTAGTTGTTGGGGGGATT 808
         ||||||||||||||| | ||||||| | |||| | || ||||| |||| ||| ||
Db   337 GGAGAAGCTGCAGCTACTAGAAGCTCCTCCAAATAAGCTCTTAGTTGGTTGGGGCAGCT 281

RESULT 11
CL746481
CL746481/c
DEFINITION   5', genomic survey sequence.
             OR_BBa0083P11.f OR BBa Oryza rufipogon genomic clone OR_BBa0083P11
LOCUS   CL746481   784 bp   DNA   linear   GSS 27-JUL-2004
ACCESSION   CL746481
VERSION   CL746481.1   GI:50688740
KEYWORDS   GSS.
SOURCE   Oryza rufipogon
  ORGANISM   Oryza rufipogon
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
             Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1 (bases 1 to 784)
  AUTHORS   Kim,H.; Yu,Y.; Stum,D.; Rao,K.; Luo,M.;
             Kudrna,D.; Muller,C.; Hatfield,J.; Soderlund,C. and Wing,R.
  TITLE   OMAP Project
  JOURNAL   Unpublished (2004)
COMMENT   Contact: Rod A. Wing
             Arizona Genomics Institute
             University of Arizona
             Forbes Building Room 303, Tucson, AZ 85721-0036, USA
             Tel: 520 626 9595
             Fax: 520 621 1259
             Email: http://genome.arizona.edu
             PCR Primers
             FORWARD: TAA TAC GAC TCA CTA TAG GG
             BACKWARD: CAC TCA TTA GGC ACC CCA
             Plate: 0083   row: P   column: 11
             Seq primer: TAA TAC GAC TCA CTA TAG GG
             Class: BAC ends.

FEATURES   Location/Qualifiers

FEATURES   Location/Qualifiers
     source   1..784
                     /organism="Oryza rufipogon"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4529"
                     /clone="OR_BBa0083P11"
                     /tissue_type="young leaves"
                     /clone_lib="OR_BBa"
                     /lab_host="DH10B-T1 phage resistant"
                     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match
Best Local Similarity   77.9%;   Score 105.2; DB 9; Length 784;
Matches 159; Conservative 0; Mismatches 28; Indels 17; Gaps 2;

Qy   585 TCTGTTTGGGGAGGCTTAAAATTATGAAAAAGCAGCTGCTGGTGAGAAGCTGAGATTACA 644
         | ||| |||||||||||| |||| ||||||||||||| |||||||||||| |||| ||
Db   204 TGTTTAGGGGAGGCTTAAGATTCTGAGAAGCAGAGATGCTGAGAAGCTAGGCTGGGAGAAT 145

Qy   645 GAAGAAATTTGA------------GTTCTACGGTTCAATTCTCCAGATTCTACAAATTACA 689
         |||||| ||| | ||| | ||| |||| | |||||||||||| |||||
Db   144 GAAGAAGCTGAAAAAACCCAACTTCTCGGCCTTCAGTTCATTTCCAGATTCTACAACTACA 85

Qy   690 GATTCTTATAATTAGGTAAAAAGCTGGACTGTTT--GGGAGCTTCGTGTCAGCCGGGAGAT 747
         |||||||| ||| ||| |||||||||||| ||||| ||||||||||||| | ||||
Db   84 GATTCTTAGAATCTGGGATAAAAAGCTGGACTTGTTTAAAGAAGAGCTTCGTGTACAACTGGGAGAT 25

Qy   748 TCTGTGAGAAGCTGCAGCGTGCTAG 771
         ||| | |||||||||||||| |||
Db   24 TCTAGGAGAAGCTGCAGCTGCTAG 1

RESULT 12
CL807668
CL807668
DEFINITION   3', genomic survey sequence.
             OR_CBa0021N19.r OR_CBa Oryza rufipogon genomic clone OR_CBa0021N19
LOCUS   CL807668   707 bp   DNA   linear   GSS 09-AUG-2004
ACCESSION   CL807668
VERSION   CL807668.1   GI:51045250
KEYWORDS   GSS.
SOURCE   Oryza rufipogon
  ORGANISM   Oryza rufipogon
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
             Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1 (bases 1 to 707)
  AUTHORS   Kim,H.; Yu,Y.; Wissotzki,M.; Yost,D.; Stum,D.; Rao,K.; Luo,M.;
             Jetty,R.; Kudrna,D.; Muller,C.; Hatfield,J.; Soderlund,C. and
             Wing,R.
  TITLE   OMAP project
  JOURNAL   Unpublished (2004)
COMMENT   Contact: Rod A. Wing
             Arizona Genomics Institute
             University of Arizona
             Forbes Building Room 303, Tucson, AZ 85721-0036, USA
             Tel: 520 626 9595
             Fax: 520 621 1259
             Email: http://genome.arizona.edu
             PCR Primers
             FORWARD: TAA TAC GAC TCA CTA TAG GG
             BACKWARD: CAC TCA TTA GGC ACC CCA
             Plate: 0021   row: N   column: 19
             Seq primer: CAC TCA TTA GGC ACC CCA
             Class: BAC ends.

FEATURES   Location/Qualifiers

RESULT 13

CL843536

LOCUS　　　　　CL843536　　　　757 bp　　　DNA　　　linear　　　GSS 09-AUG-2004
DEFINITION　　OR_CBa0072P12.r OR CBa Oryza rufipogon genomic clone OR_CBa0072P12
　　　　　　　　3', genomic survey sequence.
ACCESSION　　 CL843536
VERSION　　　 CL843536.1　GI:51089146
KEYWORDS　　　GSS.
SOURCE　　　　Oryza rufipogon
　ORGANISM　　Oryza rufipogon
　　　　　　　　Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
　　　　　　　　Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
　　　　　　　　Ehrhartoideae; Oryzeae; Oryza.
REFERENCE　　 1　(bases 1 to 757)
　AUTHORS　　　Kim,H., Yu,Y., Wissotski,M., Yost,D., Stum,D., Rao,K., Luo,M.,
　　　　　　　　Jetty,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and
　　　　　　　　Wing,R.
　TITLE　　　　OMAP project
　JOURNAL　　　Unpublished (2004)
COMMENT　　　　Contact: Rod A. Wing
　　　　　　　　Arizona Genomics Institute
　　　　　　　　University of Arizona
　　　　　　　　Forbes Building Room 303, Tucson, AZ 85721-0036, USA
　　　　　　　　Tel: 520 626 9595
　　　　　　　　Fax: 520 621 1259
　　　　　　　　Email: http://genome.arizona.edu
　　　　　　　　PCR Primers
　　　　　　　　FORWARD: TAA TAC GAC TCA CTA TAG GG
　　　　　　　　BACKWARD: CAC TCA TTA GGC ACC CCA
　　　　　　　　Plate: 0072　row: P　column: 12
　　　　　　　　Seq primer: CAC TCA TTA GGC ACC CCA
　　　　　　　　Class: BAC ends.
FEATURES　　　　　　　　 Location/Qualifiers
　　source　　　　　　　　1..757
　　　　　　　　　　　　　　/organism="Oryza rufipogon"
　　　　　　　　　　　　　　/mol_type="genomic DNA"
　　　　　　　　　　　　　　/db_xref="taxon:4529"
　　　　　　　　　　　　　　/clone="OR_CBa0072P12"
　　　　　　　　　　　　　　/tissue_type="young leaves"

ORIGIN

Query Match　　7.3%;　Score 104.6;　DB 9;　Length 757;
Best Local Similarity　72.5%;　Pred. No. 3.4e-15;
Matches 169;　Conservative　0;　Mismatches　49;　Indels　15;　Gaps　2;

Qy　　579　TAGTCGTCTCGTTTGGGAGCTTAAAATTATGAAAAAGCAGCTGCTAAGAAGCTAGCTGGTGA　638
　　　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　296　TAGGCCTGTTTAGGGGAACTTAAAAATTCTGAGAAACAGCTGCTAAGAAGCTAGCTGGTGA　355

Qy　　639　GAATCTGAAGAAATTTGA------------GTTCTACGTTCATTCTCCAGAGATTCTACA　684
　　　　　　　||||||||||||||||||　　　　　　　|||||||||||||||||||||||||||||||
Db　　356　GAATTTGGAGAAGCTGTAAACCCAGCTTCTGGCTTCATTTTTCAGATTCTACAA　415

Qy　　685　TTACAGAATTCTTATAATTTAGGTAAAAAGCTGGACTGTTTG-GGAGCTTCTCGTCAGCCGG　743
　　　　　　　|||||||||||||||||||||||||||||||||||||||　||||||||||||||||||||
Db　　416　TTACATATTCTCAGAATCTAGATAAAAAGCTAGACTGTTTGAGGAGGCTTTTAACAAATAA　475

Qy　　744　AGATTCTGTGAGAAGCTCGAGCTTCCCAAACAGACCCCTAGT　796
　　　　　　　|||||||||||||||||||||||||||||||||||||||||
Db　　476　AAATTCTAAAAGAACCTGCAGCGTCCTAGAGAGCTGTTTAACAACATTACCCCGTGGT　528

RESULT 14

AQ328957/c

LOCUS　　　　　AQ328957　　　　590 bp　　　DNA　　　linear　　　GSS 08-JAN-1999
DEFINITION　　nbxb00441l9f CUGI Rice BAC Library Oryza sativa (japonica
　　　　　　　　cultivar-group) genomic clone nbxb00441l9f, genomic survey
　　　　　　　　sequence.
ACCESSION　　 AQ328957
VERSION　　　 AQ328957.1　GI:4120807
KEYWORDS　　　GSS.
SOURCE　　　　Oryza sativa (japonica cultivar-group)
　ORGANISM　　Oryza sativa (japonica cultivar-group)
　　　　　　　　Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
　　　　　　　　Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
　　　　　　　　Ehrhartoideae; Oryzeae; Oryza.
REFERENCE　　 1　(bases 1 to 590)
　AUTHORS　　　Wing,R.A. and Dean,R.A.
　TITLE　　　　A BAC End Sequencing Framework to Sequence the Rice Genome
　JOURNAL　　　Unpublished (1998)
COMMENT　　　　Contact: Wing RA
　　　　　　　　Clemson University Genomics Institute
　　　　　　　　Clemson University
　　　　　　　　100 Jordan Hall, Clemson, SC 29634, USA
　　　　　　　　Tel: 864 656 7288
　　　　　　　　Fax: 864 656 4293
　　　　　　　　Email: rwing@clemson.edu
　　　　　　　　Seq primer: TAATACGACTCACTATAGGG
　　　　　　　　Class: BAC ends
　　　　　　　　High quality sequence stop: 324.
FEATURES　　　　　　　　 Location/Qualifiers
　　source　　　　　　　　1..590
　　　　　　　　　　　　　　/organism="Oryza sativa (japonica cultivar-group)"
　　　　　　　　　　　　　　/mol_type="genomic DNA"
　　　　　　　　　　　　　　/cultivar="Nipponbare"
　　　　　　　　　　　　　　/db_xref="taxon:39947"
　　　　　　　　　　　　　　/clone="nbxb00441l9f"
　　　　　　　　　　　　　　/tissue_type="Leaf"
　　　　　　　　　　　　　　/lab_host="E. coli DH10B"
　　　　　　　　　　　　　　/clone_lib="CUGI Rice BAC Library"
　　　　　　　　　　　　　　/note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
　　　　　　　　　　　　　　HindIII; Rice is one of two most popular grains in the
　　　　　　　　　　　　　　world. Half of the world population especially those
　　　　　　　　　　　　　　inhabiting highly populated areas of the humid tropics
　　　　　　　　　　　　　　and subtropics, rely on rice as their primary source of
　　　　　　　　　　　　　　carbohydrate. Monocotyledonous rice is a diploid plant
　　　　　　　　　　　　　　(2n=24) with a haploid genome equivalent of 431 Mbp

ORIGIN

Query Match　　7.3%;　Score 104.6;　DB 9;　Length 757;
Best Local Similarity　72.5%;　Pred. No. 3.4e-15;
Matches 169;　Conservative　0;　Mismatches　49;　Indels　15;　Gaps　2;

Qy　　579　TAGTCGTCTCGTTTGGGAGCTTAAAATTATGAAAAAGCAGCTGCTAAGAAGCTAGCTGGTGA　638
　　　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　296　TAGGCCTGTTTAGGGGAACTTAAAAATTCTGAGAAACAGCTGCTAAGAAGCTAGCTGGTGA　355

Qy　　639　GAATCTGAAGAAATTTGA------------GTTCTACGTTCATTCTCCAGAGATTCTACA　684
　　　　　　　||||||||||||||||||　　　　　　　|||||||||||||||||||||||||||||||
Db　　356　GAATTTGGAGAAGCTGTAAACCCAGCTTCTGGCTTCATTTTTCAGATTCTACAA　415

Qy　　685　TTACAGAATTCTTATAATTTAGGTAAAAAGCTGGACTGTTTG-GGAGCTTCTCGTCAGCCGG　743
　　　　　　　|||||||||||||||||||||||||||||||||||||||　||||||||||||||||||||
Db　　416　TTACATATTCTCAGAATCTAGATAAAAAGCTAGACTGTTTGAGGAGGCTTTTAACAAATAA　475

Qy　　744　AGATTCTGTGAGAAGCTCGAGCTTCCCAAACAGACCCCTAGT　796
　　　　　　　|||||||||||||||||||||||||||||||||||||||||
Db　　476　AAATTCTAAAAGAACCTGCAGCGTCCTAGAGAGCTGTTTAACAACATTACCCCGTGGT　528

/dev_stage="2 week old seedlings"
/lab_host="DH10B T1 phage resistant"
/clone_lib="OR_CBa"
/note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII;
drk treated 36 hrs before harvest"

RESULT 15

LOCUS   CG206879   431 bp   DNA   linear   GSS 21-AUG-2003
DEFINITION TOS0501 TAMU Rice Japonica Nipponbare BAC Library (Hind III) Oryza
            sativa (japonica cultivar-group) genomic clone TOSJNBh024F14h,
            genomic survey sequence.
ACCESSION   CG206879
VERSION   CG206879.1 GI:34097940
KEYWORDS   GSS.
SOURCE   Oryza sativa (japonica cultivar-group)
ORGANISM   Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1 (bases 1 to 431)
AUTHORS   Li,Y., Wu,C., Santos,T., Uhm,T., Liu,D. and Zhang,H.-B.
TITLE   BAC end sequences to close the gaps of a rice physical map at TAMU
JOURNAL   Unpublished (2003)
COMMENT   Contact: Wu C
            Department of Soil and Crop Sciences and Institute for Plant
            Genomics and Biotechnology
            Texas A & M University
            TAMU 2474, College Station, TX 77843-2474, USA
            Tel: 979 862 4800
            Fax: 979 862 4790
            Email: c-wu@neo.tamu.edu
            Seq primer: M13 Universal Reverse AACAGCTATGACCATG
            Class: BAC ends
            High quality sequence stop: 431.
FEATURES   Location/Qualifiers
     source   1..431
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="TOSJNBh024F14h"
                     /tissue_type="leaf"
                     /lab_host="E. coli DH10B"
                     /clone_lib="TAMU Rice Japonica Nipponbare BAC Library
                     (Hind III)"
                     /note="vector: V41"
ORIGIN

(Arumuganathan and Earle, 1991). The relatively small
genome of rice, three times larger than that of
Arabidopsis, makes it suitable for genomic studies. In
order to facilitate positional cloning, physical mapping
and genome sequencing of rice, we have constructed a BAC
library from Oryza sativa, Nipponbare variety. The
library contains 36,864 clones with an average insert size
of 128.5 Kb providing 10.9 haploid genome equivalents. The
deep coverage allows the isolation a particular sequence
with a probability of 99.9 %. Two high density filters,
each containing 18,432 clones (doubly spotted), represent
the whole library for colony screening."

Query Match   7.2%;   Score 104; DB 8; Length 590;
Best Local Similarity 76.3%; Pred. No. 4.6e-15;
Matches 161; Conservative 0; Mismatches 31; Indels 19; Gaps 2;

Qy   592 GGGAGCTTAAAATTATGAAAAAGCAGCAGCTAGCTGGTGAGAGAAGCTGAAGAAT 651
         ||||| ||||| ||| ||||| ||||||||||||||||||||||||||||| |||||
Db   553 GGGAGTTTAAAGATTCTAAGAAAGCAGCTGCTGAGGAAGCTAGCTGGTGAGAAATCTA 500

Qy   652 TTGAGTTCTACGTTCATTCTCCAGATTCTACAATTACAGATTCTTATAATTTAGGTAAAA 711
         |||||||| ||| |||| ||||||| ||||| ||||||||||||| ||| |||||
Db   499 --------GTTCATTTCTGATTCTACACGATTCTCAGATTCTCGGATAAAA 451

Qy   712 AGCTGGGACTGTTT--GGGAGCTTCGTCAGCCCGGAGAGTTCGTGAGGAAGCTGCAGCTGCTT 769
         ||||| |||||||| ||||||||| |||| ||||||| |||| |||||||| ||||| |
Db   450 ATCTGGACTGTTTGGGGGAGCTTCTGGCAACTGGAGAGATTTTAGGAGAAGCTGCCAGCTANT 391

Qy   770 AGAAGCTTCCCCAAACAGACCCCTAGTTGTA 800
         | ||||| |||||||||| ||||| ||| |
Db   390 AAAAGCTCCCCAAACAGGCCCCTGATTAGGA 360

ORIGIN

Query Match   7.0%;   Score 100.6; DB 9; Length 431;
Best Local Similarity 79.1%; Pred. No. 3.1e-14;
Matches 159; Conservative 0; Mismatches 34; Indels 8; Gaps 3;

Qy   598 TTAAAAATTATGAAAAGCAGCAGCTAGCTGGTGAGGAAGCTACTGGGTGAGAGAATCTGAAGAATTGAGT 657
         |||||||| ||||| |||||||| | ||| ||||||||||||| || ||| ||| |||||| || |||
Db   1 TTAAAAATTCTGAGAAGCAGCAGTTGCT-AGAAGCTAGCTGGTGAGGAAGCTAGACCTAGCTT 59

Qy   658 TCTACG-----TTCATTCTCCCAGAATTCTACAATTACAGATTCTTATATATTTAATTTAGGTAAAAA 712
         ||||| ||||| | ||| | || ||| || |||||| || | | | | ||| | ||| || ||
Db   60 TTGACTTCTCTAATTCATTTTCTACAACTGCACAGATTCTTACAACTTCTACAGATTCTTAAAATCTGGGTTAAGAA 119

Qy   713 GCTGGACTGTTTG--GGAGCTTCGTCAGCGCGGAGAGAGTTCGTGAGAAGCTGCAGCGGCTA 770
         ||||| ||||||| |||| ||||| ||| | ||||| |||| ||| ||||||| |||| ||
Db   120 GCTGAACTGTTTGGAGGAGGCTTCGACAACTGAAGAGATTCTAAGAGAAGCTGCCAGCTGTGTTA 179

Qy   771 GAAGCTTCCCCAAACAGACC 791
         ||| | || ||||||| ||
Db   180 GAAACACCTCTAAACAGGCCC 200

Search completed: September 23, 2005, 19:20:35
Job time : 3430 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: September 23, 2005, 17:01:48 ; Search time 714 Seconds
                                            (without alignments)
                                            13447.233 Million cell updates/sec

Title: US-10-602-166-2
Perfect score: 1436
Sequence: 1 cattcagaatcatcctccagc..............tatcaggtggtcaacctgg 1436

Scoring table: IDENTITY NUC
               Gapop 10.0 , Gapext 1.0

Searched: 7400704 seqs, 3343079526 residues

Total number of hits satisfying chosen parameters: 14801408

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Database :   Published Applications NA:*
               1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
               2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
               3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
               4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
               5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
               6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
               7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
               8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
               9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
               10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
               11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
               12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
               13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
               14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
               15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
               16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
               17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
               18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
               19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
               20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
               21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
               22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
               23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
               24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
               25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
               26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1436	100.0	1436	9	US-09-848-696-2	Sequence 2, Appli
2	1436	100.0	1436	18	US-10-602-166-2	Sequence 2, Appli
3	1436	100.0	3003	9	US-09-848-696-1	Sequence 1, Appli
4	1436	100.0	3003	18	US-10-602-166-1	Sequence 1, Appli
5	1196	83.3	1196	9	US-09-848-696-3	Sequence 3, Appli
6	1196	83.3	1196	18	US-10-602-166-3	Sequence 3, Appli
7	240	16.7	240	9	US-09-848-696-10	Sequence 10, Appl

8	240	16.7	240	18	US-10-602-166-10	Sequence 10, Appl
9	107.4	7.5	712	19	US-10-437-963-87679	Sequence 87679, A
10	95	6.6	1008	9	US-09-848-696-7	Sequence 7, Appli
11	95	6.6	1008	18	US-10-602-166-7	Sequence 7, Appli
12	93.4	6.5	1054	18	US-10-425-114-30221	Sequence 30221, A
13	93.4	6.5	1250	19	US-10-437-963-72585	Sequence 72585, A
14	90.4	6.3	2000	18	US-10-260-238-2616	Sequence 2616, Ap
15	77.2	5.4	17953	17	US-10-260-238-1892	Sequence 1892, Ap
16	69.2	4.8	442	19	US-10-437-963-34782	Sequence 34782, A
17	63.4	4.4	5631	19	US-10-437-963-44842	Sequence 44842, A
18	62.2	4.3	451	19	US-10-437-963-89687	Sequence 89687, A
19	61.8	4.3	5682	19	US-10-437-963-44845	Sequence 44845, A
20	58.8	4.1	2000	17	US-10-260-238-1892	Sequence 1892, Ap
21	57.6	4.0	11460	18	US-10-415-058-7	Sequence 7, Appli
22	57.6	4.0	49600	18	US-10-459-262A-4	Sequence 4, Appli
23	57.6	4.0	69300	18	US-10-415-058-6	Sequence 6, Appli
24	57.6	4.0	91552	18	US-10-415-058-5	Sequence 5, Appli
25	56.2	3.9	183	19	US-10-437-963-90519	Sequence 90519, A
26	56.2	3.9	3673778	16	US-10-312-841-1	Sequence 1, Appli
27	55.8	3.9	2000	17	US-10-260-238-1653	Sequence 1653, Ap
28	55.4	3.9	400	19	US-10-437-963-80401	Sequence 80401, A
29	55.4	3.9	835	19	US-10-437-963-97252	Sequence 97252, A
30	51	3.6	395	20	US-10-425-115-45803	Sequence 45803, A
31	51	3.6	603	19	US-10-437-963-44654	Sequence 44654, A
32	50.8	3.5	6239	17	US-10-221-613-28	Sequence 28, Appl
33	50.6	3.5	11534	19	US-10-311-455-316	Sequence 316, App
34	50.4	3.5	6735	15	US-10-311-455-1347	Sequence 1347, Ap
35	50.4	3.5	8056	20	US-10-473-126-386	Sequence 386, App
36	50.2	3.5	438	18	US-10-424-599-125688	Sequence 125688,
37	49.8	3.5	914	19	US-10-767-795-6577	Sequence 6577, Ap
38	49.8	3.5	12763	15	US-10-311-455-276	Sequence 276, App
39	49.6	3.5	5518	15	US-10-473-126-255	Sequence 255, App
40	49.6	3.5	17131	15	US-10-311-455-1025	Sequence 1025, Ap
41	49	3.4	11049	15	US-10-311-455-641	Sequence 641, App
42	49	3.4	11049	15	US-10-311-455-641	Sequence 641, App
43	48.8	3.4	8056	20	US-10-240-589C-27	Sequence 27, Appl
44	48.8	3.4	3673778	16	US-10-473-126-240	Sequence 240, App
45	48.4	3.4	456	20	US-10-312-841-2	Sequence 2, Appli
	48.4	3.4	456	20	US-10-425-115-132592	Sequence 132592,

ALIGNMENTS

RESULT 1
US-09-848-696-2
Sequence 2, Application US/0948696
Patent. No. US2001005171713A1
GENERAL INFORMATION:
; APPLICANT: AN, GYN HEUNG
; APPLICANT: JEON, JONG-SEONG
; APPLICANT: CHUNG, YONG-YOON
; APPLICANT: LEE, SI CHUL
; TITLE OF INVENTION: DNA COMPRISING RICE ANTHER-SPECIFIC GENE AND TRANSGENIC PLANT TR
; TITLE OF INVENTION: THEREWITH
; FILE REFERENCE: S-30723A
; CURRENT APPLICATION NUMBER: US/09/848,696
; CURRENT FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: KR 98-46973
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: KR 98-50126
; PRIOR FILING DATE: 1998-11-19
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.2
;
; SEQ ID NO 2
; LENGTH: 1436
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-848-696-2

Query Match   100.0%;
Best Local Similarity   100.0%;
Matches 1436;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Score 1436;   DB 9;   Length 1436;
Pred. No. 0;

RESULT 2
US-10-602-166-2
; Sequence 2, Application US/10602166
; Publication No. US20040060084A1
; GENERAL INFORMATION:
; APPLICANT: AN, GYN HEUNG
; APPLICANT: JEON, JONG-SEONG
; APPLICANT: CHUNG, YONG-YOON
; APPLICANT: LEE, SI CHUL
; TITLE OF INVENTION: DNA COMPRISING RICE ANTHER-SPECIFIC GENE AND TRANSGENIC PLANT TR
; TITLE OF INVENTION: THEREWITH
; FILE REFERENCE: 8-30723A
; CURRENT APPLICATION NUMBER: US/10/602,166
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: US/09/848,696
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: KR 98-46973
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: KR 98-50126
; PRIOR FILING DATE: 1998-11-19
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.2
; SEQ ID NO 2
; LENGTH: 1436
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-602-166-2

Query Match   100.0%;   Score 1436;   DB 18;   Length 1436;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1436;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy    1 CATTCAGAATCATCTCCAGCCTACAAATGTACTCTCCCATAATACAAGTGTCTCTATGA 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 CATTCAGAATCATCTCCAGCCTACAAATGTACTCTCCCATAATACAAGTGTCTCTATGA 60

Qy   61 TTCAAAAATTGTCCTACAATATAAACATTTCCAGCATGAAATCCATACATTAATTTCAG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCAAAAATTGTCCTACAATATAAACATTTCCAGCATGAAATCCATACATTAATTTCAG 120

Qy  121 CTAATCAGATGCTTGGGAGGGAAAAATCTAAGCGAATCAATATATGCAAAAATTGATCACTGA 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CTAATCAGATGCTTGGGAGGGAAAAATCTAAGCGAATCAATATATGCAAAAATTGATCACTGA 180

Qy  181 AGTAACTGAAAAGAGAATATCTCGTTTAACATTAGTGCTAGTATTTAATTAAACAACTAAA 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 AGTAACTGAAAAGAGAATATCTCGTTTAACATTAGTGCTAGTATTTAATTAAACAACTAAA 240

Qy  241 AAATTGTTTATATTTAGTACAAATCGAGTAGTAGCAGTAGCTAGCGGTAAAGATCG 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 AAATTGTTTATATTTAGTACAAATCGAGTAGTAGCAGTAGCTAGCGGTAAAGATCG 300

Qy  301 TGTTCCGATCACCTGAGAAACCGTCAGGTGGTTTGTCTGTGCCGTCCAGCCGATCAGAAT 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 TGTTCCGATCACCTGAGAAACCGTCAGGTGGTTTGTCTGTGCCGTCCAGCCGATCAGAAT 360

Qy  361 TCGGAGATCCGCCGTCGTTTCTTTCCTGAAAATCTGCAAGTCCCAGCGCAGCAGCGAG 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 TCGGAGATCCGCCGTCGTTTCTTTCCTGAAAATCTGCAAGTCCCAGCGCAGCAGCGAG 420

Qy  421 AGCAAGAGACCAATGGCGTGCAGGGAGTTGATACTTGACTAGCTAGCTAGCAGCAGCGCG 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 AGCAAGAGACCAATGGCGTGCAGGGAGTTGATACTTGACTAGCTAGCTAGCAGCAGCGCG 480

Qy  481 TTCGTTCCCATGCGTCGCTCTCACGCGGGCGTGGGAAGTGCCATGATCGCATGCATCGCC 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 TTCGTTCCCATGCGTCGCTCTCACGCGGGCGTGGGAAGTGCCATGATCGCATGCATCGCC 540

Qy  541 AAGATTATATTCCTCACATTTTTTCTTCCTATGCGCTCCTAGTCGTCGTTTGGGAGCTTA 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 AAGATTATATTCCTCACATTTTTTCTTCCTATGCGCTCCTAGTCGTCGTTTGGGAGCTTA 600

Qy  601 AAAATTATGAAAAACAGCTGCTGAGAAGCTGGTGGAGAATCGAGAATTTGAGTTCT 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 AAAATTATGAAAAACAGCTGCTGAGAAGCTGGTGGAGAATCGAGAATTTGAGTTCT 660

Qy  661 ACGTTCATTCTCCAGATTCTACAATTACAGATTCCTATAATTTAGGTAAAAAGCTGGACT 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 ACGTTCATTCTCCAGATTCTACAATTACAGATTCCTATAATTTAGGTAAAAAGCTGGACT 720

Qy  721 GTTTGGGGAGCTTCTGTCAGCCCGGAGATTCGTGTGAGAAGCTGCTAGAAGCTTCCC 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 GTTTGGGGAGCTTCTGTCAGCCCGGAGATTCGTGTGAGAAGCTGCTAGAAGCTTCCC 780

Qy  781 CAAACAGAGACCCCTAGTTGTTACTCTHAGCCTGATCGATTCTATTTATATATACACCTTGC 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 CAAACAGAGACCCCTAGTTGTTACTCTHAGCCTGATCGATTCTATTTATATATACACCTTGC 840

Qy  841 TCTCTAGCTTATCAAAACGTAGCCAAGACTTGAATTTAAAGCTTAAATTGAITTGAITGT 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 TCTCTAGCTTATCAAAACGTAGCCAAGACTTGAATTTAAAGCTTAAATTGAITTGAITGT 900

Qy  901 TCTTTTCATCGTAATTCACTTACCGACCTTAGTTGAATTTTTTCACGTAGTGTAAAAGTTTTA 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TCTTTTCATCGTAATTCACTTACCGACCTTAGTTGAATTTTTTCACGTAGTGTAAAAGTTTTA 960

Qy  961 TTAGAGCTGATTTGATTTATTTATTTCAGCGGGGAGTAAGCATTAGTGGTTATATAAGTAAAAGTTTTA 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 TTAGAGCTGATTTGATTTATTTATTTCAGCGGGGAGTAAGCATTAGTGGTTATATAAGTAAAAGTTTTA 1020

Qy 1021 CCTATAAATTATTAATTAATTTTCAGCCGGAGTAAGCAATTAGTGTTATGGGTTATAATCATCTGG 1080
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 CCTATAAATTATTAATTAATTTTCAGCCGGAGTAAGCAATTAGTGTTATGGGTTATAATCATCTGG 1080

Qy 1081 TATGCTTAAATCTCTTTACTTGGACAATTCGTAAATGCATTCTCGTGCCA 1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 TATGCTTAAATCTCTTTACTTGGACAATTCGTAAATGCATTCTCGTGCCA 1140
```

RESULT 3
US-09-848-696-1
; Sequence 1, Application US/09848696
; Patent No. US20010051713A1
; GENERAL INFORMATION:
; APPLICANT: AN, GYN HEUNG
; APPLICANT: JEON, JONG-SEONG
; APPLICANT: CHUNG, YONG-YOON
; APPLICANT: LEE, SI CHUL
; TITLE OF INVENTION: DNA COMPRISING RICE ANTHER-SPECIFIC GENE AND TRANSGENIC PLANT TR
; TITLE OF INVENTION: THEREWITH
; FILE REFERENCE: S-30723A
; CURRENT APPLICATION NUMBER: US/09/848,696
; CURRENT FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: KR 98-46973
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: KR 98-50126
; PRIOR FILING DATE: 1998-11-19
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.2
; SEQ ID NO 1
;   LENGTH: 3003
;   TYPE: DNA
;   ORGANISM: Oryza sativa
US-09-848-696-1

Query Match   100.0%; Score 1436; DB 9; Length 3003;
Best Local Similarity   100.0%;
Matches 1436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[Sequence alignment between Qy (query) and Db (database) — 100% identity over 1436 bases, positions 1–1436, shown as paired Qy/Db rows with matching bars (||||).]

```
Qy    541 AAGATTATATTCCTCACATTTTTCTTCCTATCCCTCCCTAGTCGTTCGTCTTA 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAGATTATATTCCTCACATTTTTCTTCCTATCCCTCCCTAGTCGTTCGTCTTA 600

Qy    601 AAAATTATGAAAAGCAGCTGCTGAGAAGCTAGCTGGTCGGTGGAGAATTCGAGTTCT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAAATTATGAAAAGCAGCTGCTGAGAAGCTAGCTGGTCGGTGGAGAATTCGAGTTCT 660

Qy    661 ACGTTCATTCTCCAGATTCTACAATTACAGATTCTTATAATTTAGGTAAAAAGTTGGACT 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ACGTTCATTCTCCAGATTCTACAATTACAGATTCTTATAATTTAGGTAAAAAGTTGGACT 720

Qy    721 GTTTGGGAGCTTCTGTCAGCCGGAGATTCTGTGAGAAGCTGCAGCTGCTAGAAGCTTCCC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GTTTGGGAGCTTCTGTCAGCCGGAGATTCTGTGAGAAGCTGCAGCTGCTAGAAGCTTCCC 780

Qy    781 CAAACAGACCCCTAGTTGTACTCTTAGCTGATCGAATTCACTCTATTTATATATACACCTTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CAAACAGACCCCTAGTTGTACTCTTAGCTGATCGAATTCACTCTATTTATATATACACCTTGC 840

Qy    841 TCTCTAGCTTATCAAACGTAGCCAAGACTTGAATTTTAAAGTTAAATGATTTGATGT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TCTCTAGCTTATCAAACGTAGCCAAGACTTGAATTTTAAAGTTAAATGATTTGATGT 900

Qy    901 TCTTTTCATCGTAATTCACTTACCGACCTTAGTCGGCCATTGAATTTTTAAAATTAATT 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TCTTTTCATCGTAATTCACTTACCGACCTTAGTCGGCCATTGAATTTTTAAAATTAATT 960

Qy    961 TTAGAGCTGATTTTGATTTTTTTTTTTTCAGGGGAATTATTTTCACGTAGTCTAAAAGTTTTA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TTAGAGCTGATTTTGATTTTTTTTTTTTCAGGGGAATTATTTTCACGTAGTCTAAAAGTTTTA 1020

Qy   1021 CCTATAAATTATTAATTTTCAGCGGGGGTAAGCATTAGTGTTTATATGGGTTTATATAATCATCTGG 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CCTATAAATTATTAATTTTCAGCGGGGGTAAGCATTAGTGTTTATATGGGTTTATATAATCATCTGG 1080

Qy   1081 TATGCTTAAATCTCTTACTTGGACTTAGTTGGGACAATTCGTAATGCAATTCGTCCGTGCCA 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TATGCTTAAATCTCTTACTTGGACTTAGTTGGGACAATTCGTAATGCAATTCGTCCGTGCCA 1140

Qy   1141 TCTCTTATAATACGGCCTGCTTGCTTGTATCTTGCACACAAGAAACTAGCTTCGGTGCAA 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TCTCTTATAATACGGCCTGCTTGCTTGTATCTTGCACACAAGAAACTAGCTTCGGTGCAA 1200

Qy   1201 AGTCCTCAAGGCGAACCGGCCTCCATCTTCTCCATCTGGCGTCCCATGT 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AGTCCTCAAGGCGAACCGGCCTCCATCTTCTCCATCTGGCGTCCCATGT 1260

Qy   1261 CGCCATCGGCCATGGCCTCTCATGGTGCAGAGATATCCAGCATGCTGAAAATTAACT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CGCCATCGGCCATGGCCTCTCATGGTGCAGAGATATCCAGCATGCTGAAAATTAACT 1320

Qy   1321 TGAGGCATATCATCTCATCATTGCAATTCAAACTTCTGGAGTTGTGTGCAGACATGTTTA 1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 TGAGGCATATCATCTCATCATTGCAATTCAAACTTCTGGAGTTGTGTGCAGACATGTTTA 1380

Qy   1381 GCTGATCAGTAAACGTTGCCGTGTTGAATTGGTCCTTATCCAGGTGGGTTCAACCTGG 1436
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GCTGATCAGTAAACGTTGCCGTGTTGAATTGGTCCTTATCCAGGTGGGTTCAACCTGG 1436
```

```
Query Match       100.0%;   Score 1436;   DB 18;   Length 3003;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1436;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy      1 CATTCAGAATCATCTCCAGCCTACAAATGTACTCTCTCCCATAATACACAAGTGTCTCTATGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CATTCAGAATCATCTCCAGCCTACAAATGTACTCTCTCCCATAATACACAAGTGTCTCTATGA 60

Qy     61 TTCAAAATTTGTCCTACAATATAAACATTTCCACCATGAAAATCCATACATTAATTTTCAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCAAAATTTGTCCTACAATATAAACATTTCCACCATGAAAATCCATACATTAATTTTCAG 120

Qy    121 CTAATCAGATGCTTGGAGGGGAAAATCTAAGCCGATTCAATATGCAAAAATTGATTCACTGA 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTAATCAGATGCTTGGAGGGGAAAATCTAAGCCGATTCAATATGCAAAAATTGATTCACTGA 180

Qy    181 AGTAACTGAAAGAGAGAATATCTCGTTTTAACATTAGTTGCTAGTATTATTAAAACAACTAAA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AGTAACTGAAAGAGAGAATATCTCGTTTTAACATTAGTTGCTAGTATTATTAAAACAACTAAA 240

Qy    241 AAAATGTTTATATTTAGTACAAAATCGACTAGTAGCAGTAGCAGCTAGCGTAAGATCG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AAAATGTTTATATTTAGTACAAAATCGACTAGTAGCAGTAGCAGCTAGCGTAAGATCG 300

Qy    301 TGTTCCGATCACCTGAGAAAACCGTGCAGGTGGTTTGTCTGTGCCCTCCAGCCGGATCAGAAT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TGTTCCGATCACCTGAGAAAACCGTGCAGGTGGTTTGTCTGTGCCCTCCAGCCGGATCAGAAT 360

Qy    361 TCGGAGAATCCGCCGTGTCGTTTCTTCTTCCTGAAATCGCAAGTCCCAGCAGCAGCAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TCGGAGAATCCGCCGTGTCGTTTCTTCTTCCTGAAATCGCAAGTCCCAGCAGCAGCAG 420

Qy    421 AGCAAGAGCAAATGGCGTGCAGAGGAGTTGATACTTTGATGCACTAGCTACTACTAGTCG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AGCAAGAGCAAATGGCGTGCAGAGGAGTTGATACTTTGATGCACTAGCTACTACTAGTCG 480

Qy    481 TTCGTTCCATGTCGCTCTCACGCGCGGTGCGAAATGTGCCATGATCCTGCAGCATCATCGCC 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TTCGTTCCATGTCGCTCTCACGCGCGGTGCGAAATGTGCCATGATCCTGCAGCATCATCGCC 540

Qy    541 AAGATTATATTCCTCACATTTTTCTTCCTATCCCTAGTCGTTCGTCTTGTTGTTTGGGAGCTTA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAGATTATATTCCTCACATTTTTCTTCCTATCCCTAGTCGTTCGTCTTGTTGTTTGGGAGCTTA 600

Qy    601 AAAATTATGAAAAGCAGCTGCTGAGAAGCTAGCTGGTCGGTGGAGAATTCGAGTTCT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAAATTATGAAAAGCAGCTGCTGAGAAGCTAGCTGGTCGGTGGAGAATTCGAGTTCT 660

Qy    661 ACGTTCATTCTCCAGATTCTACAATTACAGATTCTTATAATTTAGGTAAAAAGTTGGACT 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ACGTTCATTCTCCAGATTCTACAATTACAGATTCTTATAATTTAGGTAAAAAGTTGGACT 720

Qy    721 GTTTGGGAGCTTCTGTCAGCCGGAGATTCTGTGAGAAGCTGCAGCTGCTAGAAGCTTCCC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GTTTGGGAGCTTCTGTCAGCCGGAGATTCTGTGAGAAGCTGCAGCTGCTAGAAGCTTCCC 780

Qy    781 CAAACAGACCCCTAGTTGTACTCTTAGCTGATCGAATTCACTCTATTTATATATACACCTTGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CAAACAGACCCCTAGTTGTACTCTTAGCTGATCGAATTCACTCTATTTATATATACACCTTGC 840
```

; APPLICANT: AN, GYN HEUNG
; APPLICANT: JEON, JONG-SEONG
; APPLICANT: CHUNG, YONG-YOON
; APPLICANT: LEE, SI CHUL

RESULT 5
US-09-848-696-3
; Sequence 3, Application US/09848696
; Patent No. US2001005171713A1
; GENERAL INFORMATION:
; APPLICANT: AN, GYN HEUNG
; APPLICANT: JEON, JONG-SEONG
; APPLICANT: CHUNG, YONG-YOON
; APPLICANT: LEE, SI CHUL
; PRIOR APPLICATION NUMBER: KR 98-50126
; PRIOR APPLICATION NUMBER: KR 98-46973
; PRIOR FILING DATE: 1998-11-19
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.2
; SEQ ID NO 3
; LENGTH: 1196
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-848-696-3
; CURRENT APPLICATION NUMBER: US/09/848,696
; CURRENT FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: KR 98-46973
; PRIOR APPLICATION NUMBER: KR 98-50126
; PRIOR FILING DATE: 1998-11-03
; PRIOR FILING DATE: 1998-11-19
; NUMBER OF SEQ ID NOS: 10
; TITLE OF INVENTION: DNA COMPRISING RICE ANTHER-SPECIFIC GENE AND TRANSGENIC PLANT TRA
; FILE REFERENCE: S-30723A
; TITLE OF INVENTION: THEREWITH
; CURRENT APPLICATION NUMBER: US/09/848,696

Query Match   83.3%;   Score 1196; DB 9;   Length 1196;
Best Local Similarity   100.0%;   Pred. No. 5e-281;
Matches 1196;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

(Sequence alignment between Query (Qy, SEQ ID NO 3) and Database (Db, US-09-848-696-3), positions 1–1436/1196, full identity over the aligned region.)

RESULT 6
US-10-602-166-3
; Sequence 3, Application US/10602166
; Publication No. US20040060084A1
; GENERAL INFORMATION:
; APPLICANT: AN, GYN HEUNG
; APPLICANT: JEON, JONG-SEONG
; APPLICANT: CHUNG, YONG-YOON
; APPLICANT: LEE, SI CHUL
; TITLE OF INVENTION: DNA COMPRISING RICE ANTHER-SPECIFIC GENE AND TRANSGENIC PLANT TRA
; TITLE OF INVENTION: THEREWITH
; FILE REFERENCE: S-30723A
; CURRENT APPLICATION NUMBER: US/10/602,166
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: US/09/848,696
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: KR 98-46973
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: KR 98-50126
; PRIOR FILING DATE: 1998-11-19
; NUMBER OF SEQ ID NOS:10
; SOFTWARE: Patentin Ver. 2.2
; SEQ ID NO 3
; LENGTH: 1196
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-602-166-3

Query Match   83.3%; Score 1196; DB 18; Length 1196;
Best Local Similarity   100.0%; Pred. No. 5e-281;
Matches 1196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 CATTCAGAATCATCTCCAGCCTACACAATGTACTCTCTCCCCATAATACAAGTGTCTCTATGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CATTCAGAATCATCTCCAGCCTACACAATGTACTCTCTCCCCATAATACAAGTGTCTCTATGA 60

Qy   61 TTCAAAATTTGTCCTACATATAAAACATTTCCAGCATGAAATCCATACATTAATTTCAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCAAAATTTGTCCTACATATAAAACATTTCCAGCATGAAATCCATACATTAATTTCAG 120

Qy   121 CTAATCAGATGCTTGGACGGGGAAAAATCTAAGCGCATTCAATAATGCAAAAATTGATCACTGA 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CTAATCAGATGCTTGGACGGGGAAAAATCTAAGCGCATTCAATAATGCAAAAATTGATCACTGA 180

Qy   181 AGTAACTGAAAGAGAGAATATCTCGTTTTAACATTAGTGCTAGTAGTATTTATTAAACAACTAAA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AGTAACTGAAAGAGAGAATATCTCGTTTTAACATTAGTGCTAGTAGTATTTATTAAACAACTAAA 240

Qy   241 AAATTGTTTTATATTTTAGTACAAAATCGAGTAGTAGCAGTAGGCGTAAGATCG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 AAATTGTTTTATATTTTAGTACAAAATCGAGTAGTAGCAGTAGGCGTAAGATCG 300

Qy   301 TGTTCCGATCACCTGAGAAACCGTCAGGTGGTTTGTCTCTGTGCCGTCCAGCCGATCAGAAT 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TGTTCCGATCACCTGAGAAACCGTCAGGTGGTTTGTCTCTGTGCCGTCCAGCCGATCAGAAT 360

Qy   361 TCCGGAGATCCGCCGTCGTTTCTTTCCTGAAAATCTGCAAGTCCCAGCAGCAGCAGCAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TCCGGAGATCCGCCGTCGTTTCTTTCCTGAAAATCTGCAAGTCCCAGCAGCAGCAGCAG 420

Qy   421 AGCAAGAGCAAATGGGCGTGCAGGGAGTTTGATCACTTTGATGCACTAGCTAGCTACTAGGGCG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AGCAAGAGCAAATGGGCGTGCAGGGAGTTTGATCACTTTGATGCACTAGCTAGCTACTAGGGCG 480

Qy   481 TTCGTTCAINTGTGCTGCTCTCACGCGTGCGAAATGTGCCATGAGTCCATGCCATCATCGCC 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TTCGTTCAINTGTGCTGCTCTCACGCGTGCGAAATGTGCCATGAGTCCATGCCATCATCGCC 540

Qy   541 AAGATAATATTCCTCACATTTTTCTTCCTATCGCTCCTAGTGTCGTTGGGGAGCTTA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 AAGATAATATTCCTCACATTTTTCTTCCTATCGCTCCTAGTGTCGTTGGGGAGCTTA 600

Qy   601 AAAATTATGAAAAAGCAGCAGCTGCTGAGGAAGCTAGCTGGTGAGGAATTTGAGTTCT 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 AAAATTATGAAAAAGCAGCAGCTGCTGAGGAAGCTAGCTGGTGAGGAATTTGAGTTCT 660

Qy   661 ACGTTCATTCTCCAGATTCTACAATTACAGATTCTTATAAATTTAGGTGAAGAATCTGGACT 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 ACGTTCATTCTCCAGATTCTACAATTACAGATTCTTATAAATTTAGGTGAAGAATCTGGACT 720

Qy   721 GTTTGGGAGCTTCGTCGTCAGCCGGAGATTCGTGTGAGGAAGCTGCTAGAAGCTTCCC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GTTTGGGAGCTTCGTCGTCAGCCGGAGATTCGTGTGAGGAAGCTGCTAGAAGCTTCCC 780

Qy   781 CAAACAGACCCCTAGTTGTACTCTAGCTGATCGATCGAATCAGTCGATGGTTTAAAAGTTTA 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CAAACAGACCCCTAGTTGTACTCTAGCTGATCGATCGAATCAGTCGATGGTTTAAAAGTTTA 840

Qy   841 TCTCTAGCTTATCAAACGTAGCCAAGACTTGAAATTTAAAGCTTAAAATTGATTTGATGT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TCTCTAGCTTATCAAACGTAGCCAAGACTTGAAATTTAAAGCTTAAAATTGATTTGATGT 900

Qy   901 TCTTTTCAATCGTAATTCACTTACCGACCTTAGTCGGCATTTGAATTTTAAAATAATTATT 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCTTTTCAATCGTAATTCACTTACCGACCTTAGTCGGCATTTGAATTTTAAAATAATTATT 960

Qy   961 TTAGAGCTGATTTTGATTTTTTTTCAGCGGAAGCTACTACGTATGTAAAAGTTTA 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TTAGAGCTGATTTTGATTTTTTTTCAGCGGAAGCTACTACGTATGTAAAAGTTTA 1020

Qy   1021 CCTATAAATTATTAAATTTCAGCGGCAGGTAAGCATTAGTCGTTGTAGTGTTATCATCTGG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CCTATAAATTATTAAATTTCAGCGGCAGGTAAGCATTAGTCGTTGTAGTGTTATCATCTGG 1080

Qy   1081 TATGCTTAAATCTCTTTACTTGGACTTAGTTGGGACAAATCGTAATGCAATTCTCGTCGTTGATGT 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TATGCTTAAATCTCTTTACTTGGACTTAGTTGGGACAAATCGTAATGCAATTCTCGTCGTTGATGT 1140

Qy   1141 TCTCTATAATACGGGCCTGCTGCTTGCTCTTGTATCTGCACACAAGAGAACTAGCTG 1196
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TCTCTATAATACGGGCCTGCTGCTTGCTCTTGTATCTGCACACAAGAGAACTAGCTG 1196

RESULT 7
US-09-848-696-10
; Sequence 10, Application US/09848696
; Patent No. US20010051713A1
; GENERAL INFORMATION:
; APPLICANT: AN, GYN HEUNG
; APPLICANT: JEON, JONG-SEONG
; APPLICANT: CHUNG, YONG-YOON
; APPLICANT: LEE, SI CHUL
; TITLE OF INVENTION: DNA COMPRISING RICE ANTHER-SPECIFIC GENE AND TRANSGENIC PLANT TRA
; TITLE OF INVENTION: THEREWITH
; FILE REFERENCE: S-30723A
; CURRENT APPLICATION NUMBER: US/09/848,696
; CURRENT FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: KR 98-46973
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: KR 98-50126
; PRIOR FILING DATE: 1998-11-19
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.2
; SEQ ID NO 10
; LENGTH: 240
; TYPE: DNA
; ORGANISM: Oryza sativa

US-09-848-696-10
;
Query Match   16.7%;   Score 240;   DB 9;   Length 240;
Best Local Similarity   100.0%;   Pred. No. 3.4e-48;
Matches 240;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;
;
Qy   1197 GCAAAGTCCTCAAGGCGGAACCGCCGCTTCCATCTTCTCCATCGTCTCCGATCTCCGATCTCC 1256
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GCAAAGTCCTCAAGGCGGAACCGCC ... 60

Qy   1257 TCGTCGCCATCGCCATCGCCATCGGCGTTCTCAGGGTGGCGAGAGATATCCAGCATGCTGAAATT 1316
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TCGTCGCCATCGCCATCGCCATCGGCGTTCTCAGGGTGGCGAGAGATATCCAGCATGCTGAAATT 120

Qy   1317 AACTTGACGCATATCATCATCATCATGGCCTCTCAGGGTGCCAGAGAGATATCCAGCTGAAATT 1376
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AACTTGACGCATATCATCATCATCATGGCCTCTCAGGGTGCCAGAGAGATATCCAGCTGAAATT 180

Qy   1377 TTTAGCTGATCAGTAAACGTTGCCGTGTTGAAATTGGTTCCTTATCAGGTGGTTCAACCTGG 1436
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TTTAGCTGATCAGTAAACGTTGCCGTGTTGAAATTGGTTCCTTATCAGGTGGTTCAACCTGG 240

RESULT 8
US-10-602-166-10
; Sequence 10, Application US/10602166
; Publication No. US20040060084A1
; GENERAL INFORMATION:
; APPLICANT: AN, GYN HEUNG
; APPLICANT: JEON, JONG-SEONG
; APPLICANT: CHUNG, YONG-YOON
; APPLICANT: LEE, SI CHUL
; TITLE OF INVENTION: DNA COMPRISING RICE ANTHER-SPECIFIC GENE AND TRANSGENIC PLANT TRA
; FILE REFERENCE: S-30723A
; CURRENT APPLICATION NUMBER: US/10/602,166
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: US/09/848,696
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: KR 98-46973
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: KR 98-50126
; PRIOR FILING DATE: 1998-11-19
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.2
; SEQ ID NO 10
; LENGTH: 240
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-602-166-10

RESULT 9

Query Match   16.7%;   Score 240;   DB 18;   Length 240;
Best Local Similarity   100.0%;   Pred. No. 3.4e-48;
Matches 240;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;
US-10-602-166-10

Qy   1197 GCAAAGTCCTCAAGGCGGAACCGCCGCTTCCATCTTCTCCATCGTCTCCGATCTCCCATGGGCGTCCC 1256
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GCAAAGTCCTCAAGGCGGAACCGCC ... 60

Qy   1257 TCGTCGCCATCGCCATCGCCATCGGCGTCTCAGGGTGGCAGAGATATCCAGCATGCTGAAAATT 1316
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TCGTCGCCATCGCCATCGCCATCGGCGTCTCAGGGTGGCAGAGATATCCAGCATGCTGAAAATT 120

Qy   1317 AACTTGACGCATATCATCATCTCATCATCACTTGCATTTCAACTTCTGGATTGTGCAGACATG 1376
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AACTTGACGCATATCATCATCTCATCATCACTTGCATTTCAACTTCTGGATTGTGCAGACATG 180

Qy   1377 TTTAGCTGATCAGTAAACGTTGCCGTGTTGAAATTGGTTCCTTATCAGGTGGTTCAACCTGG 1436
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TTTAGCTGATCAGTAAACGTTGCCGTGTTGAAATTGGTTCCTTATCAGGTGGTTCAACCTGG 240

US-10-437-963-87679
; Sequence 87679, Application US/10437963
; Publication No. US20040123434A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 87679
; LENGTH: 712
; TYPE: DNA
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_86601C.1
; ORGANISM: Oryza sativa
US-10-437-963-87679

Query Match   7.5%;   Score 107.4;   DB 19;   Length 712;
Best Local Similarity   71.4%;   Pred. No. 1.3e-15;
Matches 175;   Conservative 0;   Mismatches 56;   Indels 14;   Gaps 2;

Qy   585 TCTGTTTGGGAGCTTAAAATTATGAAAAAGACGTCGCTGAGAGAGCTAGCTGGTGAGAGAATCT 644
          |||||| ||||||| ||| ||| |||| |||| |||||||| |||||||||||| |||
Db   98 TGTTTGCGGGAGCTTAAGATTCTGAGAAGCAGCTCGCTGAGAAGCTAGCTGGTAAAAATTT 157

Qy   645 GAAGAATTTGAGTTCTAC--------------GTTCATTCTCCAGATTCTACAAATTACAGAT 692
          ||| |||||| ||||| ||||||||||| ||||| ||||| |||||
Db   158 GGAGAAGCTGGAAAAACCCAGCTTCTAGCTGTTCATTTCTCACAACTACAGAT 217

Qy   693 TCTTATAATTAGGTAAAAAAAGCTGGACTGTGTT---GGGAGCTTTCGTCTCAGCGCGGAGATTCT 750
          |||| |||| ||| ||||| ||||| || | | |||||||| | |||| |||||| |||||
Db   218 TCTTAGAATCTAGGTAAAAAGCTAAATTGTTTTGAGGGAGCTTCGTGTAACTGGAGATTCT 277

Qy   751 GTGAGAAGCTGCAGCTGCTAGAAGCTTCCCAAACACAGACCCCCAGTGTTGTACTTGTACTTCAGCTGA 810
          |||| ||||| |||| ||||||||||||| ||||| |||||| ||||| |||| |||||
Db   278 ATAAGAAGATGTAGCTGCTAGAAGCTCCCTAAACAAGCCCTTAGTATGACTTAAACATGA 337

Qy   811 TCGAT 815
          |||
Db   338 TCGTT 342

RESULT 10
US-09-848-696-7
; Sequence 7, Application US/09848696
; Patent No. US20010051713A1
; GENERAL INFORMATION:
; APPLICANT: AN, GYN HEUNG
; APPLICANT: JEON, JONG-SEONG
; APPLICANT: CHUNG, YONG-YOON
; APPLICANT: LEE, SI CHUL
; TITLE OF INVENTION: DNA COMPRISING RICE ANTHER-SPECIFIC GENE AND TRANSGENIC PLANT TRA
; FILE REFERENCE: S-30723A
; CURRENT APPLICATION NUMBER: US/09/848,696
; CURRENT FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: KR 98-46973
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: KR 98-50126
; PRIOR FILING DATE: 1998-11-19
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.2
; SEQ ID NO 7
; LENGTH: 1008

; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-848-696-7

Qy   Query Match   6.6%;   Score 95;   DB 9;   Length 1008;
Db   Best Local Similarity 100.0%;   Pred. No. 1.7e-12;
     Matches 95; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1197 GCAAAGTCCTCAAGGCAGAACCGCCTCCATCTTCTCCTTCCCAGCTTCCTCCCATGGCGTCCC 1256
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GCAAAGTCCTCAAGGCAGAACCGCCTCCATCTTCTCCTTCCCAGCTTCCTCCCATGGCGTCCC 60

Qy   1257 TCGTCGCCATCGCCATGGCTCTCATGGGTG 1291
          ||||||||||||||||||||||||||||||
Db   61 TCGTCGCCATCGCCATGGCTCTCATGGGTG 95

RESULT 11
US-10-602-166-7
; Sequence 7, Application US/10602166
; Publication No. US20040060084A1
; GENERAL INFORMATION:
; APPLICANT: AN, GYN HEUNG
; APPLICANT: JEON, JONG-SEONG
; APPLICANT: CHUNG, YONG-YOON
; APPLICANT: LEE, SI CHUL
; CURRENT APPLICATION NUMBER: US/10/602,166
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: US/09/848,696
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: KR 98-46973
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: KR 98-50126
; PRIOR FILING DATE: 1998-11-19
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.2
; FILE REFERENCE: S-30723A
; TITLE OF INVENTION: DNA COMPRISING RICE ANTHER-SPECIFIC GENE AND TRANSGENIC PLANT TRA
; TITLE OF INVENTION: THEREWITH
;
; SEQ ID NO 7
; LENGTH: 1008
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-602-166-7

Qy   Query Match   6.6%;   Score 95;   DB 18;   Length 1008;
Db   Best Local Similarity 95%;   Pred. No. 1.7e-12;
     Matches 95; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1197 GCAAAGTCCTCAAGGCAGAACCGCCTCCATCTTCTCCTTCCCAGCTTCCTCCCATGGCGTCCC 1256
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GCAAAGTCCTCAAGGCGAACCGCCTCCATCTTCTCCTTCCCAGCTTCCTCCCATGGCGTCCC 60

Qy   1257 TCGTCGCCATCGCCATGGCTCTCATGGGTG 1291
          ||||||||||||||||||||||||||||||
Db   61 TCGTCGCCATCGCCATGGCTCTCATGGGTG 95

RESULT 12
US-10-425-114-30221
; Sequence 30221, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 30221
; LENGTH: 1054
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-OSROM202018E07_FLI
US-10-425-114-30221

Qy   Query Match   6.5%;   Score 93.4;   DB 18;   Length 1054;
Db   Best Local Similarity 98.9%;   Pred. No. 4.3e-12;
     Matches 94; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   1197 GCAAAGTCCTCAAGGCAGAACCGCCTCCATCTTCTCCTTCCCAGCTTCCTCCCATGGCGTCCC 1256
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GCAAAGTCCTCAAGGCGAACCGGCTCCATCTTCTCCTTCCCAGCTTCCTCCCATGGCGTCCC 60

Qy   1257 TCGTCGCCATCGCCATGGCTCTCATGGGTG 1291
          ||||||||||||||||||||||||||||||
Db   61 TCGTCGCCATCGCCATGGCTCTCATGGGTG 95

RESULT 13
US-10-437-963-72585
; Sequence 72585, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 72585
; LENGTH: 1250
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1250)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_72950C.1
US-10-437-963-72585

Qy   Query Match   6.5%;   Score 93.4;   DB 19;   Length 1250;
Db   Best Local Similarity 98.9%;   Pred. No. 4.7e-12;
     Matches 94; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   1197 GCAAAGTCCTCAAGGCAGAACCGCCTCCATCTTCTCCTTCCCAGCTTCCTCCCATGGCGTCCC 1256
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GCAAAGTCCTCAAGGCGAACCGGCTCCATCTTCTCCTTCCCAGCTTCCTCCCATGGCGTCCC 60

Qy   1257 TCGTCGCCATCGCCATGGCTCTCATGGGTG 1291
          ||||||||||||||||||||||||||||||
Db   61 TCGTCGCCATCGCCATGGCTCTCATGGGTG 95

RESULT 14
US-10-260-238-2616
; Sequence 2616, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.

Qy   Query Match   6.5%;   Score 93.4;   DB 18;   Length 1250;
Db   Best Local Similarity 98.9%;   Pred. No. 4.7e-12;
     Matches 94; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   1197 GCAAAGTCCTCAAGGCAGAACCGCCTCCATCTTCTCCTTCCCAGCTTCCTCCCATGGCGTCCC 1256
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13 GCAAAGTCCTCAAGGCGAACCGGCTCCATCTTCTCCTTCCCAGCTTCCTCCCATGGCGTCCC 72

Qy   1257 TCGTCGCCATCGCCATGGCTCTCATGGGTG 1291
          ||||||||||||||||||||||||||||||
Db   73 TCGTCGCCATCGCCATGGCTCTCATGGGTG 107

; APPLICANT: Moughamer, Todd G.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiyaki
; APPLICANT: Kreps, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 2616
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: N region
; LOCATION: (1496)..(1496)
; OTHER INFORMATION: n = any nucleotide
US-10-260-238-2616

Query Match   6.3%; Score 90.4; DB 17; Length 2000;
Best Local Similarity   71.1%; Pred. No. 3.3e-11;
Matches 155; Conservative   0; Mismatches 46; Indels 17; Gaps 2;

Qy   593 GGAGCTTAAAATTATGAAAAGCAGCTGCTGAGAAGCTAGCTGGTGAGAAATT 648
           ||||||||||||||||| |||||||||||||||||||||||||||| |||
Db   17088 GGAGCTTAAGATTCTGAGAAGCCAACTGCTAGAAGCTAGCTGGTGAGAAGC 17147

Qy   653 TGAGTTCTAC--------------GTTCATTCTCCAGATTCTACAATTACAGATTCTTA 697
           ||||||||| ||||||| |||||||||||||||||||
Db   17148 TGGAAAACCCAGCTTCTTGTCTTCTAGTTCATTTCATTTCTACATCTATAGATTCTTA 17203

Qy   698 TAATTTAGGTAAAAAGCTGGACTGTTTGG--GAGCTTCTGTCAGCCGGAGATTCTGTGAG 755
           |||||||||||| ||||||||||||||| |||||||||||||||||||||| |||
Db   17204 AGAAGTAGACTATTTGAGG----------AGCTGAAATTCTAAAGAAAACTGCAAATAC 17254

Qy   756 AAGCTGCAGCTGCTAGAAGCTTCCCCAAACAGACCCCT 793
           |||||||||||||||||||||||||||||||||||||
Db   17255 TAGAAGCTCACTCAAAACAGGGCCCTTAGT 17282

Qy   709 AAAAGCTGGACTGTTTGGGAGCTTCTGTCAGCCGGAGATTCTGTGAGAAGGTCAGCTGC 768
Db   17204 ...

Search completed: September 23, 2005, 19:36:05
Job time : 720 secs

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 17953
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-415-058-4

Query Match   5.4%; Score 77.2; DB 18; Length 17953;
Best Local Similarity   68.3%; Pred. No. 1.8e-07;
Matches 142; Conservative   0; Mismatches 53; Indels 13; Gaps 2;

Qy   589 TTTGGGAGCTTAAAATTATGAAAAGCAGCTGCTGAGAAGCTAGCTGGTGAGAAATCTGAAG 648
           ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   17088 TTGGGGAGCTTAAGATTCTGAGAAGTAGCTGGTGGTTGGTTAGCTTAGCTTCTGATAATCTGGAA 17147

Qy   649 AATTTGAGTTCTACGTTCATTCTCCAGATTCTACAATTACAGATTCTTA 708
           ||||||||||||||||||| ||||||| |||||||||||||||||||||
Db   17148 AAGTTGGGTTT----TTCAACTTCTGGCTTCTATAACTACACAGATTCTTA 17203

Qy   709 AAAAGCTGGACTGTTTGGGAGCTTCTGTCAGCCGGAGATTCTGTGAGAAGGTCAGCTGC 768
           ||| ||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   17204 AGCTTGAAAATTCTAAAGAAAACTGCAAATAC 17254

Qy   769 TAGAAGCTTCCCCAAACAGACCCCTAGT 796
           ||||||||||||| |||||||| ||||
Db   17255 TAGAAGCTCACTCAAAACAGGGCCCTTAGT 17282

RESULT 15
US-10-415-058-4
; Sequence 4, Application US/10415058
; Publication No. US20040060081A1
; GENERAL INFORMATION:
; APPLICANT: Wisconsin Alumni Research Foundation
; APPLICANT: United States Department of Agriculture
; APPLICANT: Leong, Sally A.
; APPLICANT: Farman, Mark L.
; APPLICANT: Chauhan, Rajinder
; APPLICANT: Durfee, Timothy J.
; TITLE OF INVENTION: Plant Gene That Confers Resistance To Strains Of Magnaporthe Gris
; FILE REFERENCE: Warf-0145
; CURRENT APPLICATION NUMBER: US/10/415,058
; CURRENT FILING DATE: 2003-04-11
; TITLE OF INVENTION: Having AVR C039 Cultivar Specificity Gene
; PRIOR APPLICATION NUMBER: USSN 60/242,313
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: USSN 60/303,897
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 14

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2005, 16:39:48 ; Search time 197 Seconds
                                                (without alignments)
                                                11927.377 Million cell updates/sec

Title: US-10-602-166-2
Perfect score: 1436
Sequence: 1 cattcagaatcatctccagc........tatcaggtggtcaacctgg 1436

Scoring table: IDENTITY_NUC
                   Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                   Maximum Match 100%
                   Listing first 45 summaries

Database :   Issued_Patents_NA:*
                   1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
                   2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
                   3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
                   4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
                   5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
                   6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

•   Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	46.2	3.2	7218	1	US-08-232-463-14	Sequence 14, Appl
2	44.8	3.1	13158	2	US-08-687-080-105	Sequence 105, App
3	44.6	3.1	832	4	US-09-621-976-2813	Sequence 2813, Ap
4	43.6	3.0	1664976	4	US-08-916-421B-1	Sequence 1, Appli
5	43.4	3.0	1664976	4	US-09-692-570-1	Sequence 1, Appli
6	43.4	3.0	1039	4	US-09-902-540-1280	Sequence 1280, Ap
7	42.4	3.0	612	1	US-09-902-540-1357	Sequence 1357, Ap
8	42.4	3.0	740	3	US-08-998-416-563	Sequence 563, App
9	41.8	2.9	570	1	US-09-621-976-1972	Sequence 1972, Ap
10	41.8	2.9	2748	1	US-09-949-016-1448	Sequence 1448, Ap
11	41.8	2.9	36154	4	US-09-621-976-13190	Sequence 13190, A
12	40.6	2.8	601	1	US-09-949-016-202462	Sequence 202462, A
13	40.6	2.8	6243	2	US-09-056-075-1	Sequence 1, Appli
14	40.4	2.8	212139	2	US-09-949-016-87645	Sequence 87645, A
15	40	2.8	601	1	US-09-949-016-16065	Sequence 16065, A
16	40	2.8	19124	2	US-09-949-016-12505	Sequence 12505, A
17	40	2.8	670689	4	US-09-949-016-14207	Sequence 14207, A
18	40	2.8	670690	4	US-09-270-767-3490	Sequence 3490, Ap
19	39.8	2.8	897	4	US-09-270-767-18772	Sequence 18772, A
20	39.8	2.8	897	4	US-09-949-016-122354	Sequence 122354, A
21	39.6	2.8	601	4	US-08-487-826B-13	Sequence 13, Appl
22	39.6	2.8	19124	2	US-09-949-016-16935	Sequence 16935, A
23	39.6	2.8	36618	4	US-09-949-016-12550	Sequence 12550, A
24	39.6	2.8	45842	4	US-09-949-016-17327	Sequence 17327, A
25	39.6	2.8	45842	4	US-09-949-016-15189	Sequence 15189, A
26	39.6	2.8	74730	4	US-09-949-016-91213	Sequence 91213, A
27	39.4	2.7	601	4	US-09-949-016-91214	Sequence 91214, A
28	39.4	2.7	601	4	US-09-949-016-166398	Sequence 166398, A
29	39.4	2.7	601	4	US-09-949-016-17422	Sequence 17422, A
30	39.4	2.7	164061	4	US-09-949-016-16423	Sequence 16423, A
31	39.4	2.7	167708	4	US-09-949-016-14353	Sequence 14353, A
32	39.4	2.7	175236	4	US-09-949-016-11808	Sequence 11808, A
33	39.4	2.7	636591	4	US-09-949-016-13388	Sequence 13388, A
34	39.4	2.7	636591	4	US-09-949-016-122355	Sequence 122355, A
35	39.2	2.7	601	6	US-09-949-016-15092	Sequence 15092, A
36	39.2	2.7	3095	6	US-09-949-016-202463	Sequence 202463, A
37	39.2	2.7	3095	5	US-08-749-522-5	Sequence 5, Appli
38	39.2	2.7	133719	3	US-08-467-504-3	Sequence 3, Appli
39	39	2.7	3974	3	US-09-949-016-12528	Sequence 12528, A
40	39	2.7	2317	3	US-09-949-016-12386	Sequence 12386, A
41	39	2.7	601	3	US-09-949-016-16915	Sequence 16915, A
42	38.8	2.7	76810	4	US-09-790-988-1	Sequence 1, Appli
43	38.8	2.7	263693	4		
44	38.8	2.7	263694	4		
45	38.8	2.7	640681	4		

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No 5670367
GENERAL INFORMATION:
;   APPLICANT: DORNER, F.
;   APPLICANT: SCHEIFLINGER, F.
;   APPLICANT: FALKNER, F. G.
;   TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
;   NUMBER OF SEQUENCES: 52
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Foley & Lardner
;   STREET: 1800 Diagonal Road, Suite 500
;   CITY: Alexandria
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22313-0299
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/232,463
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: 435
;   FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US/07/935,313
;   FILING DATE:
;   APPLICATION NUMBER: EP 91 114 300.6
;   FILING DATE: 26-AUG-1991
;   ATTORNEY/AGENT INFORMATION:
;   NAME: BENT, Stephen A.
;   REGISTRATION NUMBER: 29,768
;   REFERENCE/DOCKET NUMBER: 30472/114 IMMU
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (703)836-9300
;   TELEFAX: (703)683-4109
;   TELEX: 899149
;   INFORMATION FOR SEQ ID NO: 14:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 7218 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   IMMEDIATE SOURCE:
;   CLONE: pTzgpt-P1s
US-08-232-463-14

Query Match   3.2%;   Score 46.2;   DB 1;   Length 7218;

```
Best Local Similarity   6.3%; Pred. No. 0.033;
Matches  27; Conservative 218; Mismatches 186; Indels   0; Gaps   0;

Qy     899 GTTCTTTTCATCGTAATTCACTTACCGACCTTAGTCGCCGCATTTGAATTTTAAAAATAAT 958
           :                                                   : ::
Db    1064 GATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1123

Qy     959 TTTTAGAGCTGATTTGATTTTTTTTTTTTTCAGCGGGAATTAATTTTCACGTAYGTAAAAGTTT 1018
           :::                                                    :::
Db    1124 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1183

Qy    1019 TACCTATAAATTATTAATTTAATTTCAGCGGAGTAAGCANTAGTGTTATGGGGTTATAATCATCT 1078
           :::: :                                            : ::: :
Db    1184 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1243

Qy    1079 GGTAYGCTTAAATCTTTTTACTTGGACTTAGTTGTGGACANATTCGTAAATGCATTCTCGTGC 1138
           ::                                                    :::
Db    1244 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1303

Qy    1139 CATCTCTATAATACGGCCCTGCCTAGCTTGCTCTTGTATCTGCACACAAGAACTAGCTGGC 1198
           :                                                      :
Db    1304 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1363

Qy    1199 AAAGTCCTCAAGGCGAACCGCCTCCATCTTCCTCCCAGCTCCTCCCATGGGCGGTCCCTC 1258
           : ::: :                                          : :: :
Db    1364 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1423

Qy    1259 GTCGCCATGCCCATGCCCTCTCATGGTGCAGAGAGATATCCAGCATGCTGAAATTAA 1318
           ::::: :                                         :  ::: :
Db    1424 YYYYYYYYYYYGTACCAAAAATTCTTCTATCTCTTTAACTACTTGCATAGATAGGTAAATTAC 1483

Qy    1319 CTTGACGCATA 1329
           ||||
Db    1484 AGTGATGCCTA 1494
```

```
RESULT 2
US-08-687-080-105
Sequence 105, Application US/08687080
Patent No. 595427
GENERAL INFORMATION:
  APPLICANT: Gregory Dolganov
  TITLE OF INVENTION: Human RAD50 Gene and Methods of Use Thereof
  NUMBER OF SEQUENCES: 175
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Dehlinger & Associates
    STREET: 350 Cambridge Avenue, Suite 250
    CITY: Palo Alto
    STATE: CA
    COUNTRY: USA
    ZIP: 94306
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/687,080
    FILING DATE: 17-JUL-1996
    CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/592,126
    FILING DATE: 26-JAN-1996
  ATTORNEY/AGENT INFORMATION:
    NAME: Sholtz, Charles K.
    REGISTRATION NUMBER: 38,615
    REFERENCE/DOCKET NUMBER: 4600-0111.30
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (415) 324-0880
    TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 105:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 13158 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: double
    TOPOLOGY: linear
  MOLECULE TYPE: DNA (genomic)
  HYPOTHETICAL: NO
  ANTI-SENSE: NO
  ORIGINAL SOURCE:
    INDIVIDUAL ISOLATE: 5' END OF INTRON 21 OF RAD50 GENOMIC
    INDIVIDUAL ISOLATE: SEQUENCE
US-08-687-080-105
```

```
Query Match             3.1%;  Score 44.8; DB 2; Length 13158;
Best Local Similarity  52.1%;  Pred. No. 0.12;
Matches 100; Conservative   0; Mismatches  92; Indels   0; Gaps   0;

Qy     920 TTACCGACCTTAGTCGCCGCATTTGAATTTTTAAAAATAATTTTGAGCTGATTTGATT 979
           ||||| |                                   : :::
Db    3008 TTAAAAACCATAGCTTTAAGTGAATTGTTTTATCTTTTCCTGAATACTAATATATCTTC 3067

Qy     980 TTTTTTCAGCGGGAATTAATTTTCACGTAYGTAAAAGTTTTACCTTATAAAATTAATTTTT 1039
           :: ::                                          : ::
Db    3068 TCTCANCTCGCATTAATTCTTTGTCTAAATATTTTAGTTTTACCTGTTGTTTTAAATCATCA 3127

Qy    1040 CAGCGGGAGTAAGCANTAGTGTTATGGGGTTATAATCATCTGGTAYGCTTAAATCTCTTAC 1099
           : ::                                           ::::
Db    3128 CAAATGAGTAATAATAATTATTCGTGGGTGGTTTCATAGCCGCAATGTTTCCCAAATTTCTTTGT 3187

Qy    1100 TTGGACTTAGTT 1111
           ||||
Db    3188 TCATTCTTGCTT 3199
```

```
RESULT 3
US-09-621-976-2813
Sequence 2813, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
  APPLICANT: Dumas Milne Edwards, J.B.
  APPLICANT: Jobert, S.
  APPLICANT: Giordano, J.Y.
  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
  FILE REFERENCE: GENSET.054PR2
  CURRENT APPLICATION NUMBER: US/09/621,976
  CURRENT FILING DATE: 2000-07-21
  NUMBER OF SEQ ID NOS: 19335
  SOFTWARE: Patent.pm
  SEQ ID NO 2813
  LENGTH: 832
  TYPE: DNA
  ORGANISM: Homo sapiens
  FEATURE:
  NAME/KEY: CDS
  LOCATION: 235..399
US-09-621-976-2813
```

```
Query Match             3.1%;  Score 44.6; DB 4; Length 832;
Best Local Similarity   9.3%;  Pred. No. 0.027;
Matches  29; Conservative 154; Mismatches 128; Indels   0; Gaps   0;

Qy     893 TTTGATGTTCTTTTCATCGTAATTCACTTACCGACCTTAGTCGCCGCATTTGAATTTTAAA 952
           :|:| :||:|                               :|:|:
Db      43 KKKKAAWWKYWKTWTWWRYWRYAWWGTYKKKAMCRTKTKKKKKKGYWMMWYWGWRRSYWAWW 102

Qy     953 AATAATTTTAGAGCTGATTTGATTTTTTTTTCAGCGGGAATTAATTTTCACGTAYGTAA 1012
           :|:| |::::: :|:|:||:|                         :|:||: :
Db     103 TRTWTGYAYYRSMMYWWRYRCWKKKAYYRKTTCYSSKGWTWWKRWKKAWTTWWWKKTYYW 162

Qy    1013 AAGTTTTACCTTATAAATTATTAATTTTTCAGCGGAGTAAGCATTAGTGTTATGGGTTTATAA 1072
           :::|:||:|                                 :|:|::|:||||:
Db     163 AATRYWWMMCWTKRWRASWWYCWWWGKARKWSTWRKSRSYASARSAXRCCYSCSWGAMSW 222

Qy    1073 TCATCTGGTAYGCTTAAATCTCTTTACTTGGACTTAGTTGTGGGACAANTTCGTAAATGCATTC 1132
           :||| :| : :                         :|:|: :|:| : :|:|
```

```
Db    223  KYMWRMWRWRGWAYGAGMKAWRASCMMRRKYAGKSKTSYKSMWMCWTRSWKYCYTKARWT 282
                  :: .::::.::::.::  :::.::::.:::::::.:::::: .: :::: .:  ::.::
Qy    1133 TCGTGCCATCTCTATANATACGGCCGTCCTCAGCTTTGCTCTTGTATCTGCACACAGAACTA 1192
Db    283  GYYCYRKGGMWGKRGRWYASKKYMWKRWWCWARMYRYSTGTRASMWWRRKYYTMMMKWW 342
                  ::.::::::::::: .: ::.:::.:::.:: ::.: .:..: :: .:: :  ::.  :: ::
Qy    1193 GCTCGCAAAGT 1203
                  :::::::::
Db    343  KYAWARAAWRW 353
```

; RESULT 4
; US-08-916-421B-1
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococ
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Version 3.1
;
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98259)..(98259)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (191989)..(191989)

; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
```

```
;   NAME/KEY: misc feature
;   LOCATION: (1084830)..(1084830)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1130881)..(1130881)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1096846)..(1096846)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1110988)..(1110988)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1119881)..(1119881)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1313224)..(1313224)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1349473)..(1349473)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1349491)..(1349491)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1603734)..(1603734)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1602912)..(1602912)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1569020)..(1569020)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1470091)..(1470091)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1377998)..(1377998)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1664854)..(1664855)
;   OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1
```

```
Query Match          3.0%;    Score 43.6;  DB 4;   Length 1664976;
Best Local Similarity 47.2%;  Pred. No. 4;
Matches 202; Conservative  0; Mismatches 219; Indels  7;  Gaps  2;

Qy    674 AGATTCTACAAATTACAGATTCTTATAATTTAGGTAAAAAGCTGGACTGTTTGGGAGCTTC 733
          ||||   ||   |||| ||||||  |  ||||| |||| |||||| ||| ||  |||||||
Db 251603 AGATTAGCCAAATGGAAAAAACCTGAACCTAAGGTTAAAAGGTTATTAGCAAGATAT 251662

Qy    734 TGTCAGCCCGGAGATTCTGTGAGAAGCTGCAGCTGCTGCTAGAAGCTTCCCCAAACACGACCCCT 793
          ||||       |  |||  |  ||||| | ||  | |||| ||||| |||||  | || |||||
Db 251663 GCTAAGCTGTAAGTTCAGCTGATGAGGAGCTGTATTAAGATAGTAAATACGAGATTT 251722

Qy    794 AGTTGTACTCTAGCTGATCGATCACTCTATATTTATATACACCTTGCTCTCTAGCTTATC 853
          ||||     |||||  | |   || |  |||| ||| ||| | ||| | | |||||
Db 251723 CTTTATGCTTATTGTTATTTTTACATAATATTTTTATTACCAATTTAT---AATTTGTC 251779

Qy    854 AAACGTAGCCAAGACTTGAATTTTAAAGCTTAAATTGAATTTGATGTTCTTTTCATCGTA 913
          | | |||||  | | ||||  |||| ||| || |  || ||| ||| ||  ||||| |||
Db 251780 GTAATACACTAGGACTAGGATTTTTAAATTTTATATGGAATTTGGGAAAGTTATCTCGTTCA 251839

Qy    914 ATTCACTTACCGGACCTTAGTCGGGCATTTGAATTTTTAAAAATAATTTTTAGAGCTGATTT 973
          |  |  | ||   | ||    |||  ||| |  |||| |||  | |||| ||||| |
Db 251840 ATACATTTATAATTAGGAAAAACCCATTTAAAAATCGATAATCATTTATTTTAACCTTTTT 251899

Qy    974 TGATTTTTTTTCAGCGGAATTTATTTTCACG----TATGTAAAGTTTTACCTATAAAT 1029
          |||||||||||||||  |  ||||||||||||||    |||  |||||| ||| ||||
Db 251900 TATCTAAATTCTAAGGGTAGCTTATTTAAAAAATTTATTATTTTTGGGATTTGTTAAATTA 251959

Qy   1030 TATTAATTTTCAGCCGGAGTAAGCATAGTGTTATGGGTTATAATCATCTGGTATGCTTAA 1089
          ||||  ||||||  |||||||||| ||||||| ||||| |||||| | ||| ||| ||||
Db 251960 TAGGGATTTTTAAAAAATTCTACTTAATTGTTTTTATTTGAGATTTCTCCAATGATTAA 252019

Qy   1090 ATCTCTTT 1097
          ||||||||
Db 252020 TTTTTATT 252027
```

```
RESULT 5
US-09-692-570-1
;   Sequence 1, Application US/09692570
;   Patent No. 6797466
;   GENERAL INFORMATION:
;   APPLICANT: Bult et al.
;   TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococc
;   FILE REFERENCE: PB275C1
;   CURRENT APPLICATION NUMBER: US/09/692,570
;   CURRENT FILING DATE: 2003-01-14
;   PRIOR APPLICATION NUMBER: US 60/024,428
;   PRIOR FILING DATE: 1996-08-22
;   PRIOR APPLICATION NUMBER: US 08/916,421
;   PRIOR FILING DATE: 1997-08-22
;   NUMBER OF SEQ ID NOS: 20
;   SOFTWARE: Patentin version 3.1
;   SEQ ID NO 1
;   LENGTH: 1664976
;   TYPE: DNA
;   ORGANISM: Methanococcus jannaschii
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (84808)..(84808)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (84812)..(84812)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (28120)..(28120)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (28222)..(28222)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (28257)..(28258)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (84773)..(84773)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (98155)..(98159)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (98239)..(98239)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (98266)..(98266)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (98343)..(98343)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (103998)..(103998)
```

~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (148948)..(148948)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (163385)..(163385)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (191989)..(191989)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (191995)..(191995)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (231980)..(231980)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (234187)..(234187)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (234220)..(234220)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (234814)..(234814)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (309398)..(309398)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (309418)..(309418)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (312837)..(312837)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (312993)..(312993)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (319226)..(319226)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (559167)..(559167)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (559241)..(559241)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (600592)..(600592)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (622708)..(622708)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (657081)..(657081)
~ OTHER INFORMATION: n equals a, t, c, or g

~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (657203)..(657203)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (674435)..(674435)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (682442)..(682442)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (713652)..(713652)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (741684)..(741684)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (779455)..(779455)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (779676)..(779676)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (855539)..(855539)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (871619)..(871619)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (1084830)..(1084830)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (1096846)..(1096846)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (1119881)..(1119881)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (1313224)..(1313224)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (1310988)..(1310988)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (1130881)..(1130881)
~ OTHER INFORMATION: n equals a, t, c, or g
~ FEATURE:
~ NAME/KEY: misc_feature
~ LOCATION: (1349473)..(1349473)
~ OTHER INFORMATION: n equals a, t, c, or g

Query Match
Best Local Similarity   3.0%;   Score 43.6; DB 4; Length 1664976;
Matches 202; Conservative 47.2%;   Pred. No. 4;
            0; Mismatches 219;   Indels 7;   Gaps 2;

Qy   674 AGAATTCTACAAATTACAGATTCTTATATAATTTAGGTAAAAAAGCTGGACTGTTGGGAGCTTC 733
           ||||| ||| || | |||| ||||||||||||||||||||||||||||||
Db 251603 AGATTAGCCAAATGGAAAAAACCTGAACCTRAGGTTAAAAAAGGTTATTTAGCAAGATAT 251662

```
Qy      734  TGTCAGCCCGAGATTCTGTGTGAGGAAGCTGCAGCGTCGCTAGAAGAAGCTTCCCAAAACAGACCCCCT  793
             ||| |  ||||||      |||     |     | ||||||||||
Db   251663  GCTAAGCCTGTAAGTTCAGCTGATCGAGGGAGCTGTATTAAGAGATGATTAAATAGAGAATTT  251722

Qy      794  AGTTGTACTCTAGCTGATCGGAATTCACTCTATTATATACACCTTGCCTCTCTAGCTTATC  853
             ||  ||| |||||||||||||
Db   251723  CTTTATGCTTATTGTATTTTTACATAATAATTTTATTTTTATTCCAATTTAT---AATTTGTC  251779

Qy     1096  TTACTTGGGACTTAGTTGGGACAATTCGTAANTGCAATCTCGT  1136
             ||||||||
Db      854  AAACGTAGCCAAGACTGAATTAAAGCTAAATGATGAGTTAGATTTGATGTCTTTTCANCGGTA  913

Qy     1030  TATTAAATTTCAGCGGGAGTAAGCATTAGTGTTATGGGTTAATCATCTGGTATGCTTAA  1089
             ||||   |||||
Db   251780  GTAATACACTAGGACTAGGATAGTTTTAATTTATATGATTGGGAAAGTTATCTGGTTCA  251839

Qy      914  ATTCACTTACCGACCTTAGTCGGCATTGAATTTTTAAAATAATTAATTTTTGAGTGTGATTT  973
             ||||                |||                     |||
Db   251840  ATACANTTATAATTAGGAAAACCCATTTAAAATCTGATATCATTATTTTTAACCTTTTT  251899

Qy      974  TGATTTTTTTTTTCAGCGGGAATTTATTTTCACG----TATGTAAAAGTTTTACCTATATAAAT  1029
             ||| ||||||||||||||                |||
Db   251900  TATCTAAATTCTAAGGGTACCTTATTTTAAAAAATTTATTTTTATTTTGTTGTTAAAATTA  251959

Qy     1030  TATTAAATTTCAGCGGGAGTAAGCATTAGTGTTATGGGTTAATCATCTGGTATGCTTAA  1089
             ||||   |||||
Db   251960  TAGGGATTTTAAAAAATTCTACTTAAATGTTTTTTTTTATTTTGAGATTTCTCCAATGATTAA  252019

Qy     1090  ATCCTTTT  1097
             ||  |||
Db   252020  TTTTTATT  252027
```

```
Query Match      3.0%;  Score 43.4;  DB 4;  Length 1039;
Best Local Similarity  45.5%;  Pred. No. 0.068;
Matches 15;  Conservative  0;  Mismatches 186;  Indels  0;  Gaps  0;
```

```
Qy      796  TTGTACTCTAGCTGATCGATTCACTCTATTATATACACCTTGCTCTCTAGCTTATCAA  855
             ||||| |||   |||
Db      195  TTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTTTTTTTTTTTTTTTTGTTT  136

Qy      375  TTGTTTTTTTGTTATTAATTTAATTTTTATTGATTCATTTGATTTAATTTTTTTT  316
             |||||||||||||
Db     1036  TTTTCAGCGGAGTAAGCATTAGTGTTATGGGTTATAATCATCTGGTATGCTTAAATCTCT  1095

Qy      856  ACGTAGCCAAGACTTGAATTTAAAGCTAAATTGATTGATGTTCATTTCATCGTAAT  915
             |||||||
Db      135  TTATTTTTTTTTATTAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTGTTTG  76

Qy      916  TCACTTACCGACCTTAGTCGGCATTGAATTTTAAAAATAATTTTTTAGAGCTGATTTTG  975
             ||||||||
Db     1096  TTACTTGGGACTTAGTTGGGACAATTCGTAANTGCAATCTCGT  1136

Qy      255  TGTTTTTTTATATTTTTTTTATATTTTTTAATTTTTTTTTTTTT  196
             ||  |||
Db       75  TTTTTAGTTTTTTGAAGGGCTTATTTTTTTTATTTTTGT  35

Qy      976  ATTTTTTTTTCAGCGGGGAATTAATTTCACGTATGTAGTATAANTAATTATTAA  1035
             |||||||||
```

```
Query Match      3.0%;  Score 42.4;  DB 4;  Length 612;
Best Local Similarity  48.3%;  Pred. No. 0.097;  Mismatches 123;  Indels  0;  Gaps  0;
Matches 115;  Conservative  0;
```

```
Qy      874  TTTTAAAGCTTAAATTGATTTTGATGTTCTTTTCATCGTAANTCACTTACCGGACCTTAGT  933
             |||                     |||
Db      437  TTTTTTTTTTTTTTTTTTTTTTTATTTTTTTTTTTTTTTTTTTTTTTT  378

Qy      934  CGGGCATTGAATTTTAAAAAATAATTTTAGAGCTGCGATTTTGATTTTTTTCAGCGCGGA  993
             |||||| ||||||         |||||||
Db      377  TTTTTTTTATTTTTTTATTTTTTTTTTTTTTTTTTANTTTTTTTTTTTAT  318

Qy      994  TTTTATTTCACGTATGTAAAAGTTTTACCTATAAATTATTAAAATTTTCAGCCGGAGTAAGCA  1053
             |||||||||                    |||
Db      317  TTTTANTTTTTTTTTTTTAATTTTTTTTTTTTTTTTTTATTTTTATT  258

Qy     1054  TTAGTGTTATGGGTTAATCATCTGGTATGCTTAAAATCTCTTACTTGGACTTAGTT  1111
             |||||||
Db      257  TTTTTTTTTATTTTTTTTTTTTTTTTTTTATTTTTTTTTTTTTTATT  200
```

; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: NO. 6239264artis Corporation
;   STREET: 3054 Cornwallis Road
;   CITY: Research Triangle Park
;   STATE: No. 6239264th Carolina
;   COUNTRY: USA
;   ZIP: 27709
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/998,416
;   FILING DATE: 24-DEC-1997
;   CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: CH 0016/97
;   FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
;   NAME: Meigs, J. Timothy
;   REGISTRATION NUMBER: 38,241
;   REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC11976
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 919-541-8587
;   TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 563:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 740 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
;   ORGANISM: PAG1390RP
; US-08-998-416-563

  Query Match   3.0%; Score 42.4; DB 3; Length 740;
  Best Local Similarity   49.1%; Pred. No. 0.11;
  Matches 112; Conservative   0; Mismatches 116; Indels   0; Gaps   0;

Qy   814 ATTCACTCTATTTATATACACCTTGCTCTCTAGCTTATCAAACGTAGCCAAGACTTGAA 873
          ||| ||| ||||| |||
Db   106 ATTTATTGATTTTGTTGACATTGTGATAATGTTATTATAAAAATATAATTTAA 165

Qy   874 TTTTAAAGCTTAAAATGATTTTGAGTGTCTTTTCATCGTAATTCACTTACCGACCTTAGT 933
          |||||
Db   166 TATTATTTTTATATAATATTATTATATCTCAGTCATTGATAAATATGAATTATAT 225

Qy   934 CGGCATTTGAATTTTAAAAATAATTAATTTTAGAGCTGATTTTGATTTTTTTCAGCGGGAA 993
          ||||| ||||
Db   226 TCCATTATTAATTGTTTAGGATAAACATAAAATTAATAATAATAAACTTATTATTTAAGTTCAA 285

Qy   994 TTTTATTTTCACGTATGTAAAAGTTTTACCTATAAAGTTTATTAATAATTAATTTCA 1041
          |||| ||||
Db   286 TAAATATGTTCATATTTATATGATTAATCATAAGCTATTCGATATAA 333

RESULT 9
US-09-621-976-1972
; Sequence 1972, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;   APPLICANT: Dumas Milne Edwards, J.B.
;   APPLICANT: Jobert, S.
;   APPLICANT: Giordano, J.Y.
;   TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;   FILE REFERENCE: GENSET.054PR2
;   CURRENT APPLICATION NUMBER: US/09/621,976
;   CURRENT FILING DATE: 2000-07-21
;   NUMBER OF SEQ ID NOS: 19335
;   SOFTWARE: Patent.pm
;   SEQ ID NO 1972

;   LENGTH: 570
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 92..340
;   NAME/KEY: Big_peptide
;   LOCATION: 92..325
;   OTHER INFORMATION: Von Heijne matrix
;   OTHER INFORMATION: score 3.70000004768372
;   OTHER INFORMATION: seq NLITAPFSSPSTS/TP
; US-09-621-976-1972

  Query Match   2.9%; Score 41.8; DB 4; Length 570;
  Best Local Similarity   61.5%; Pred. No. 0.14;
  Matches 67; Conservative   0; Mismatches 42; Indels   0; Gaps   0;

Qy   940 TTGAATTTTTAAAAATAATTAATTTTTAGAGCTGATTTTGATTTTTTTTCAGCGGGAATTTATT 999
          |||||| |||||| ||||||||
Db   390 TTGAAATTTTTTTATAAATATGTGAAATTTCTTAGGATTTTTTTACCAGCCTTTTGTTGTTTACA 449

Qy   1000 TTCACGTATGTAAAAGTTTTACCTATATAAAATTATTAATTAATTTTCACAGAAT 1048
          |||| ||||||
Db   450 GACCCAAAATGTAAATATTAAAAAATAAATATTATTGCAATTTTCTACAGAAT 498

RESULT 10
US-09-949-016-1448
; Sequence 1448, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;   APPLICANT: VENTER, J. Craig et al.
;   TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;   TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;   FILE REFERENCE: CL001307
;   CURRENT APPLICATION NUMBER: US/09/949,016
;   CURRENT FILING DATE: 2000-04-14
;   PRIOR APPLICATION NUMBER: 60/241,755
;   PRIOR FILING DATE: 2000-10-14
;   PRIOR APPLICATION NUMBER: 60/237,768
;   PRIOR FILING DATE: 2000-10-20
;   PRIOR APPLICATION NUMBER: 60/231,498
;   PRIOR FILING DATE: 2000-10-03
;   PRIOR APPLICATION NUMBER: 60/231,498
;   PRIOR FILING DATE: 2000-09-08
;   NUMBER OF SEQ ID NOS: 207012
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 1448
;   LENGTH: 2748
;   TYPE: DNA
;   ORGANISM: Human
; US-09-949-016-1448

  Query Match   2.9%; Score 41.8; DB 4; Length 2748;
  Best Local Similarity   61.5%; Pred. No. 0.34;
  Matches 67; Conservative   0; Mismatches 42; Indels   0; Gaps   0;

Qy   940 TTGAATTTTTAAAAATAATTAATTTTTAGAGCTGATTTTGATTTTTTTTCAGCGGGAGT 999
          |||||| |||||| ||||||
Db   2582 TTGAAAATTTTTTATAAATATGTATGGAAATTCCTTAGGATTTTTTTACCAGGCGGGAATTTACA 2641

Qy   1000 TTCACGTATGTAAAAGTTTTACCTATAAAGTTTATTAATTATTTCAGCGCGGAAGT 1048
          |||| ||||||
Db   2642 GACCCAAAATGTAAATATTAAAAAATAAATAATATTTTTGCAATTTCTACACAGAAT 2690

RESULT 11
US-09-949-016-13190
; Sequence 13190, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;   APPLICANT: VENTER, J. Craig et al.
;   TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;   TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;   FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 13190
; LENGTH: 36154
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13190

Query Match   2.9%; Score 41.8; DB 4; Length 36154;
Best Local Similarity   61.5%; Pred. No.1.5;
Matches   67; Conservative   0; Mismatches   42; Indels   0; Gaps   0;

```
Qy    940 TTGAATTTTTAAAAATAATATAATTTTTAGAGCTGGATTGATTTTTGATTTTTTTTATT 999
          ||||| ||| ||| ||||| ||||||
Db  33988 TTGAAATTTTTATAAATATGTATGGGAAATTCTTAGGATTTTTTTGTTTGTTTACA 34047

Qy   1000 TTCACGTATGTAAAAGTTTTACCTTATAAAATTAAATTTCAGCGGGAGT 1048
Db  34048 GACCCAAAATGTAAATATTAAAAATAAATATTTGCAATTTTCTACAGAAT 34096
```

RESULT 12
US-09-949-016-202462/c
; Sequence 202462, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 202462
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-202462

Query Match   2.8%; Score 40.6; DB 4; Length 601;
Best Local Similarity   48.4%; Pred. No. 0.31;
Matches 109; Conservative   1; Mismatches 115; Indels   0; Gaps   0;

```
Qy     40 ATAATACAAGTGTCTCTATGATTCAAAAATTGTCCTACAATATAAACATTCCAGCCATGA 99
Db    347 ATACAAAAAGCATCTTGCATATATTAGGTCGTCCTCACTAAAATTGTAACYAGCGATCATTATTA 288

Qy    100 AATCCATACACATTAATTTTCAGCTAATCAGATGCTTGGAGGGAAAAATCTAAGCGGATTCAA 159
Db    287 GATACAAAAATAAAACATTTGAGTACAGGGATTAGCAGTTTACAAAATAATTTTATGTATA 228

Qy    160 TATGCAAAAATTGATCACTGAAGTAACTGAAAGAGAATATCTCGTTTTAACATTAGTGCT 219
Db    227 AGAGTATATGGTAATTTTATGAAATTATTTAAAATAAGTGTATTATTATATAAATTAGGTGT 168

Qy    220 AGTATTTATTAAACAACTAAAAAAAATTGTTTATATTTTA 264
Db    167 AAATTTCAGAAAAAAAGTAAAAATACAGGGAAATCTCTAAGCACACA 123
```

RESULT 13
US-09-056-075-1
; Sequence 1, Application US/09056075
; Patent No. 5955368
; GENERAL INFORMATION:
; APPLICANT: Johnson, Eric A.
; APPLICANT: Bradshaw, Marite
; APPLICANT: Rood, Julian
; TITLE OF INVENTION: Expression System for Clostridium
; TITLE OF INVENTION: Species
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/056,075
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95238
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6243 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3770..4013
; OTHER INFORMATION: /note= "RP4 origin of DNA transfer (oriT) from
; OTHER INFORMATION: plasmid RP4"
US-09-056-075-1

Query Match   2.8%; Score 40.4; DB 2; Length 6243;
Best Local Similarity   49.1%; Pred. No. 1.3;
Matches 107; Conservative   0; Mismatches 111; Indels   0; Gaps   0;

```
Qy     40 ATAATACAAGTGTCTCTATGATTCAAAAATTGTCCTACAATATAAACATTCCAGCCATGA 99
Db   1227 ATAAAAAATAAAAAATATAAAAAAATAAAAAAATCTAAGCGGATTCAA 1286

Qy    100 AATCCATACACATTAATTTTCAGCTAATCAGATGCTTGGAGGGAAAAATCTAAGCGGATTCAA 159
Db   1287 AAAAAATAAAAAATATAAAAAAATAAAAAATATAAAAATAAAAA 1346

Qy    160 TATGCAAAAATTGATCACTGAAGTAACTGAAAGAGAATATCTCGTTTTAACATTAGTGCT 219
Db   1347 AATAAAAATAAAATAAAAAAAAATAAAAAAATAAAAAATATAAAAAAATA 1406

Qy    220 AGTATTTATTAAACAACTAAAAAAAATTGTTTATATTTTA 257
Db   1407 TTTTTTATTTAAAGTTTGAAAAAAAATTTTTTTTATATTA 1444
```

RESULT 14
US-09-949-016-16065

```
; Sequence 16065, Application US/0994016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16065
; LENGTH: 212139
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16065

Query Match                    2.8%;  Score 40.4;  DB 4;  Length 212139;
Best Local Similarity   47.9%;  Pred. No. 10;  Mismatches 126;  Indels   0;  Gaps   0;
Matches  116;  Conservative   0;  Mismatches 126;  Indels   0;  Gaps   0;

Qy     810 ATCGATTCACTCTATTTTATATACACCTTGCTCTCTAGCTTATCAAACGTAGCCAAGACT 869
               |||| |||||| ||||  ||||| | || |||||||| ||| | ||| |||| ||||
Db   51215 ACCCACCCATCCTCTTGTATACGTTAAATCATCCTCTAGCTTACTTATAATACCTTAATACA 51274

Qy     870 TGAATTTAAAGCTTAAATTGATTTGATGTTCTTTTCATCGTAATTCACTTACCGACCT 929
                ||||| |||||||||||| ||||||| ||| |||| ||||||||| |||| |
Db   51275 ATACAAAATCGTGTGTAAATAGCTATATATATTGATTTTTATTGTATTATTTTATTGTTI 51334

Qy     930 TAGTCGGCATTTGAATTTTTTAAAAATAATAATTTTTAGAGCTGATTTTGATTTTTTTCAGC 989
               ||| ||||||||||| |||||||  |||  ||||| ||||||||| || |||||||||
Db   51335 TTGTTGTTTTTTAAATTTTTTCCTGAATATTATGTGTGTTATTGAGATGTTACTTT 51394

Qy     990 GGAATTTATTTTCACGTAGTGATAAAAGTTTTACCTATAAATTATTAAATTTCAGCGGGAGTA 1049
               ||||||  |||||| |  ||||| ||| |||||||||||||  ||||||||| ||
Db   51395 GTGATTTAAGTAATTCAAGTAAAACTCGTTAAGCATTTATTTCTGTTTTTAGTGTAGCATA 51454

Qy    1050 AG 1051
               ||
Db   51455 AG 51456
```

RESULT 15
```
US-09-949-016-87645
; Sequence 87645, Application US/0994016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 87645
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-87645

Query Match                    2.8%;  Score 40;  DB 4;  Length 601;
Best Local Similarity   49.1%;  Pred. No. 0.46;  Mismatches 110;  Indels   0;  Gaps   0;
Matches 106;  Conservative   0;  Mismatches 110;  Indels   0;  Gaps   0;

Qy     881 GCTTAAATTGATTTGATGTTCTTTTCATCGTAATTCACTTACCGACCTTAGTCGGCATT 940
               ||||| ||||| |||||| ||| |||||||||||||| ||| || ||||||||||
Db     203 GCTTAAACTTCTGCCCTGGGAATCTTTTTGCTAATCCCTGGTTATCCGTGTCTTTAAATAGAGCCT 262

Qy     941 TGAATTTTAAAAATAATTTTTTAGAGCTGATTTTGATTTTTTTTTTTCAGCGGGAATTTATTT 1000
               |||| |||||||||||| |||||||||||||| |||| |||||||||
Db     263 TTCTTATTTGATTACTTATTTATATATACTTGCCTTCGTTTTTTTTTAGACAAAATTTATT 322

Qy    1001 TCACGTAGTGTAAAAGTTTTACCTATAAATTATTAAATTTTCAGCGGGAGTAAGCATTAGTGT 1060
               ||||| |||||||| ||||||| |||||| ||||||| ||||||||||||
Db     323 TCATAATTGTTGAATTAAACAGACACAGCTTCTACACTTATGGGGTTTCCTTTTCTTTCTGGGA 382

Qy    1061 TATGGGTTATATAATCATCTGGTATGCTTAAAATCTCTT 1096
               ||||| |||||||| |||||  ||
Db     383 GAATGGGGCTCGGTGACAAATTAATGCAGAAAACTTTTT 418

Search completed: September 23, 2005, 19:24:03
Job time : 205 secs
```

This Page Blank (uspto)

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: September 23, 2005, 14:46:25 ; Search time 4312 Seconds
                (without alignments)
                16136.756 Million cell updates/sec

Title: US-10-602-166-2
Perfect score: 1436
Sequence: 1 cattcagaatcatctccagc........tatcaggtggttcaacctgg 1436
Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0
Searched: 4708233 seqs, 2422607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%   Maximum Match 100%
                 Listing first 45 summaries

Database : GenEmbl:*
        1: gb_ba:*
        2: gb_htg:*
        3: gb_in:*
        4: gb_om:*
        5: gb_ov:*
        6: gb_pat:*
        7: gb_ph:*
        8: gb_pl:*
        9: gb_pr:*
       10: gb_ro:*
       11: gb_sts:*
       12: gb_sy:*
       13: gb_un:*
       14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1436	100.0	1436	6	BD242728	BD242728 DNA conta
2	1436	100.0	3003	6	BD242727	BD242727 DNA conta
3	1303	90.7	145828	8	AP005392	AP005392 Oryza sat
4	1303	90.7	154188	8	BD242729	BD242729 DNA conta
5	1196	83.3	1196	6	BD242729	BD242729 DNA conta
6	1079.8	75.2	2757	8	AF042275	AF042275 Oryza sat
7	240	16.7	240	6	BD242735	BD242735 DNA conta
8	184.6	12.9	598	8	AK109240	AK109240 DNA conta
9	137.8	9.6	127952	8	OSJN00063	OSJN00063 Oryza sat
10	137.8	9.6	156002	8	OSJN00050	OSJN00050 Oryza sat
11	127.4	8.9	126544	8	AP005106	AP005106 Oryza sat
12	126	8.8	156649	8	AC144738	AC144738 Oryza sat
13	121.6	8.5	95209	2	AP004323	AP004323 Oryza sat
14	121.6	8.5	187273	8	AP005966	AP005966 Oryza sat
15	120	8.4	134159	2	AP003625	AP003625 Oryza sat
16	120	8.4	168253	8	AP003686	AP003686 Oryza sat
17	118.8	8.3	139421	8	AC087545	AC087545 Oryza sat
18	118.8	8.3	300029	8	AB017097	AB017097 Oryza sat
19	118	8.2	158742	8	AC009042	AC009042 Oryza sat
20	118	8.2	272997	8	AB017069	AB017069 Oryza sat
21	117.4	8.2	185095	8	AP003436	AP003436 Oryza sat
22	117.4	8.2	185481	2	AC091088	AC091088 Oryza sat
23	116.8	8.1	91053	8	AL831811	AL831811 Oryza sat
24	115.4	8.0	167239	2	CNS08CAO	AL607104 Oryza sat
25	115.4	8.0	132647	8	OSJN01012	AC147962 Oryza sat
26	114.2	8.0	300029	8	AC073393	AC073393 Oryza sat
27	114.2	7.9	141148	8	AB017100	AB017100 Oryza sat
28	113.2	7.9	141148	8	AB017100	AB017100 Oryza sat
29	112.4	7.9	200626	8	AC109929	AC109929 Oryza sat
30	112.4	7.8	129387	8	AP005161	AP005161 Oryza sat
31	112.4	7.8	138203	8	AP004183	AP004183 Oryza sat
32	112.4	7.8	186132	8	AP005528	AP005528 Oryza sat
33	111.8	7.8	158173	8	AP005439	AP005439 Oryza sat
34	110.4	7.7	141336	8	AC092778	AC092778 Oryza sat
35	110.4	7.7	181786	8	AP003762	AP003762 Oryza sat
36	110	7.7	127642	8	AP003714	AP003714 Oryza sat
37	110	7.7	138584	8	AC093952	AC093952 Oryza sat
38	109.8	7.6	149349	8	AC097175	AC097175 Oryza sat
39	109.8	7.6	162132	8	AC134236	AC134236 Oryza sat
40	109.6	7.6	116758	8	AC134237	AC134237 Oryza sat
41	108	7.5	144395	8	AP004134	AP004134 Oryza sat
42	107.4	7.5	2708	8	AC119291	AC119291 Oryza sat
43	107.4	7.5	174588	8	AK119775	AK119775 Oryza sat
44	107.4	7.5	186991	8	AP004989	AP004989 Oryza sat
45	107.2	7.5	121885	8	AP003728	AP003703 Oryza sat

ALIGNMENTS

```
RESULT 1
BD242728
LOCUS       BD242728                1436 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  DNA containing rice anther-specific gene and transgenic plants
            transformed thereby.
ACCESSION   BD242728
VERSION     BD242728.1  GI:33052498
KEYWORDS    JP 2002528125-A/2.
SOURCE      Oryza sativa
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 1436)
  AUTHORS   An,G.; Jeon,J.S.; Chung,Y.Y. and Lee,S.
  TITLE     DNA containing rice anther-specific gene and transgenic plants
            transformed thereby
  JOURNAL   Patent: JP 2002528125-A 2 03-SEP-2002;
            SYNGENTA PARTICIPATIONS AG
COMMENT     OS   Oryza sativa (rice)
            PN   JP 2002528125-A/2
            PD   03-SEP-2002
            PF   02-NOV-1999 JP 2002579761
            PR   03-NOV-1998 KR  1998/46973,19-NOV-1998 KR  1998/50126 PI
            GYNHEUNG AN,JONG SEONG JEON,YONG YOON CHUNG,SICHUL LEE PC
            C12N15/09,A01H5/00,C07K14/415,C12N15/00
            CC   DNA containing rice anther-specific gene
            and transgenic plants
            CC   transformed
            CC   thereby
            FH   Key             Location/Qualifiers
            FT   source          1..1436
            FT                   /organism='Oryza sativa (rice)'.
FEATURES             Location/Qualifiers
     source          1..1436
                     /organism="Oryza sativa"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4530"
ORIGIN

Query Match             100.0%;  Score 1436;  DB 6;  Length 1436;
Best Local Similarity   100.0%;  Pred. No. 0;
```

Matches 1436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 CATTCAGAATCATCTCCAGCCTACAAATGTACTCTCCCATAATACAAGTGTCTCTATGA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CATTCAGAATCATCTCCAGCCTACAAATGTACTCTCCCATAATACAAGTGTCTCTATGA 60

Qy     61 TTCAAAATTGTCCTACAATATAAACATTCCAGCATAATCATCATTAAATTTCAG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCAAAATTGTCCTACAATATAAACATTCCAGCATAATCATCATTAAATTTCAG 120

Qy    121 CTAATCAGATGCTTGGGACGGAAAAATCTAAGCGAATTCAATATGCAAAAAATTGATCACTGA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTAATCAGATGCTTGGGACGGAAAAATCTAAGCGAATTCAATATGCAAAAAATTGATCACTGA 180

Qy    181 AGTAACTGAAAGAGAATATCTCGTTTTAACATTAGTGCTAGTATTATTATATAAAGAATTAACT 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AGTAACTGAAAGAGAATATCTCGTTTTAACATTAGTGCTAGTATTATTATATAAAGAATTAACT 240

Qy    241 AAATGTTTATATTTAGTACAAATCGAGTAGTAGCAGTAGCACGTAAGATCG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AAATGTTTATATTTAGTACAAATCGAGTAGTAGCAGTAGCACGTAAGATCG 300

Qy    301 TGTTCCGATCACCTGAGAAAACGTCAGGTGGTTTGTTCTGTGCCCGTTCCAGCCGATCAGAAT 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TGTTCCGATCACCTGAGAAAACGTCAGGTGGTTTGTTCTGTGCCCGTTCCAGCCGATCAGAAT 360

Qy    361 TCGGGAGATCCGCCGTCGTTCTTTCCTGAAATCTGCAAGTCCCAGCAGCAGCAGCAGCAG 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TCGGGAGATCCGCCGTCGTTCTTTCCTGAAATCTGCAAGTCCCAGCAGCAGCAGCAGCAG 420

Qy    421 AGCAAGAGCAAATGGGCGTGCAGGGAGTTTGATCACTTGAGCACTAGGCG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AGCAAGAGCAAATGGGCGTGCAGGGAGTTTGATCACTTGAGCACTAGGCG 480

Qy    481 TTCGTTCCATGTCGCTCTCACGCGTGCGAATGTGCCATGCATCATCATCGCC 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TTCGTTCCATGTCGCTCTCACGCGTGCGAATGTGCCATGCATCATCATCGCC 540

Qy    541 AAGATTATATTCCTCACATTTTTCTCCTATCGCTCGTCTGTTCTGTTGGGGACTTA 600
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAGATTATATTCCTCACATTTTTCTCCTATCGCTCGTCTGTTCTGTTGGGGACTTA 600

Qy    601 AAATTATGAAAAGCAGCTGCTGAGGAATCTGAGGAATTTGAGTTCT 660
            ||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAATTATGAAAAGCAGCTGCTGAGGAATCTGAGGAATTTGAGTTCT 660

Qy    661 ACGTTCATTCTCCAGATTCTACAATTACAGATTCTTATATAATTTAGGTAAAAAGCTTGGACT 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ACGTTCATTCTCCAGATTCTACAATTACAGATTCTTATATAATTTAGGTAAAAAGCTTGGACT 720

Qy    721 GTTTGGGGAGCTTCTGTCAGCGGAGAGATTCTGAGAGTGCAGCTGCCTAGAGAGCTTCCC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GTTTGGGGAGCTTCTGTCAGCGGAGAGATTCTGAGAGTGCAGCTGCCTAGAGAGCTTCCC 780

Qy    781 CAAACAGACCCCTAGTTGTACTCTAGCTGATCAGTCACTCTATTTTATATACACCTTGC 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CAAACAGACCCCTAGTTGTACTCTAGCTGATCAGTCACTCTATTTTATATACACCTTGC 840

Qy    841 TCTCTAGCTTATCAAACGTAGCCAAGACTTGAATTTAAAGCTTAAAATTGATTTTGATGT 900
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TCTCTAGCTTATCAAACGTAGCCAAGACTTGAATTTAAAGCTTAAAATTGATTTTGATGT 900

Qy    901 TCTTTTTCATCGTAATTCACTTACCGGACCTTAGTCGGCATTTGAATTTTAAAAATAATTT 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TCTTTTTCATCGTAATTCACTTACCGGACCTTAGTCGGCATTTGAATTTTAAAAATAATTT 960

Qy    961 TTAGAGCTGATTTGATTTTTTTTTCAGCGGAGTAAGCAATTAGTTAATGTTAAAAGTTTTA 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TTAGAGCTGATTTGATTTTTTTTTCAGCGGAGTAAGCAATTAGTTAATGTTAAAAGTTTTA 1020

Qy   1021 CCTATAAAATTATTAATTTTTCAGCGGAGTAAGCAATTAGTTATGGTTATGAATCATCTCGG 1080
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CCTATAAAATTATTAATTTTTCAGCGGAGTAAGCAATTAGTTATGGTTATGAATCATCTCGG 1080

Qy   1081 TATGCTTAAATCTCTTTACTTGAGCTTAGTTGGGACTTAGTTGGGGACAATTCGTAATGCATTCTCGTAGTGGCCA 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TATGCTTAAATCTCTTTACTTGAGCTTAGTTGGGACTTAGTTGGGGACAATTCGTAATGCATTCTCGTAGTGGCCA 1140

Qy   1141 TCTCTATAATACGGCGCTGCTGAGCTTGCTCTTGTATCTGGCACAAGAGAACTAGCTGGGCAA 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TCTCTATAATACGGCGCTGCTGAGCTTGCTCTTGTATCTGGCACAAGAGAACTAGCTGGGCAA 1200

Qy   1201 AGTCCTCAAGGCGGAACCGGCCTCCACTCTTCTCCTTCCCAGCTTCCTCCTCGT 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AGTCCTCAAGGCGGAACCGGCCTCCACTCTTCTCCTTCCCAGCTTCCTCCTCGT 1260

Qy   1261 CGCCATCGCCATGCCGCCATGGCTCTCCATCATGTGCCAGAGATATCCAGCAGATTAACT 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CGCCATCGCCATGCCGCCATGGCTCTCCATCATGTGCCAGAGATATCCAGCAGATTAACT 1320

Qy   1321 TGACGGCATCATCTCATCATCGCATTCAATTCAACTTCTGGATTGTGCAGACATGTTTA 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 TGACGGCATCATCTCATCATCGCATTCAATTCAACTTCTGGATTGTGCAGACATGTTTA 1380

Qy   1381 GCTGATCAGTAAACGTTGCCGTGTTGAATTGGTTCTTATCAGGTGGTTCAACCTGG 1436
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GCTGATCAGTAAACGTTGCCGTGTTGAATTGGTTCTTATCAGGTGGTTCAACCTGG 1436
```

RESULT 2
BD242727

LOCUS   BD242727   3003 bp   DNA   linear   PAT 17-JUL-2003
DEFINITION   DNA containing rice anther-specific gene and transgenic plants
            transformed thereby.
ACCESSION   BD242727
VERSION   BD242727.1   GI:33052497
KEYWORDS   JP 2002528125-A/1.
SOURCE   Oryza sativa
   ORGANISM   Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1 (bases 1 to 3003)
   AUTHORS   An,G., Jeon,J.S., Chung,Y.Y. and Lee,S.
   TITLE   DNA containing rice anther-specific gene and transgenic plants
            transformed thereby
   JOURNAL   Patent: JP 2002528125-A 1 03-SEP-2002;
            SYNGENTA PARTICIPATIONS AG
COMMENT   OS   Oryza sativa (rice)
            PN   JP 2002528125-A/1
            PD   03-SEP-2002
            PF   02-NOV-1999 JP 2000579761
            PR   03-NOV-1998 KR   1998/46973,19-NOV-1998 KR   1998/50126 PI
            GYNHEUNG AN,JONG SEONG JEON, YONG YOON CHUNG,SICHUL LEE PC
            C12N15/09,A01H5/00,C07K14/415,C12N15/00
            CC   DNA containing rice anther-specific gene
            and transgenic plants
            CC   thereby   transformed
FEATURES   Location/Qualifiers
            source   1..3003
                       /organism="Oryza sativa"
                       /mol_type="genomic DNA"
                       /db_xref="taxon:4530"
   FH   Key   Location/Qualifiers
   FT   source   1..3003
            FT   /organism="Oryza sativa (rice)'.
ORIGIN

Query Match   100.0%;   Score 1436;   DB 6;   Length 3003;
Best Local Similarity   100.0%;
Matches 1436;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

```
Qy      1 CATTCAGAATCATCTCCAGCCTACAAATGTACTCTCCCATAATACAAGTGTCTCTATGA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CATTCAGAATCATCTCCAGCCTACAAATGTACTCTCCCATAATACAAGTGTCTCTATGA 60
```

```
Qy    61  TTCAAAATTTGTTCCTACATAAAATCCATACATTTCCAGCATGATGAAATCCATACATTAATTTCAG  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  TTCAAAATTTGTTCCTACATAAAATCCATACATTTCCAGCATGATGAAATCCATACATTAATTTCAG  120

Qy   121  CTAATCAGATGCTTGGAGGGGAAAAATCTAACGCGATTCAATATGCAAAAAATTGATCACTGA  180
          |||||||||||||||||||||||||||||||||||||||||||
Db   121  CTAATCAGATGCTTGGAGGGGAAAAATCTAACGCGATTCAATATGCAAAAAATTGATCACTGA  180

Qy   181  AGTAACTGAAAAGAGAAATATCTCCGTTTTAACATTAGTGCTAGTATTTATTAAAACACTAAA  240
          |||||||||||||||||||||||||||||||||||||||||||
Db   181  AGTAACTGAAAAGAGAAATATCTCCGTTTTAACATTAGTGCTAGTATTTATTAAAACACTAAA  240

Qy   241  AAAATTGTTTATATTTTAGTACAAAATGAGTAGCAGTAGGAGCTAGGGTAAGATCG  300
          |||||||||||||||||||||||||||||||||||||||||||
Db   241  AAAATTGTTTATATTTTAGTACAAAATGAGTAGCAGTAGGAGCTAGGGTAAGATCG  300

Qy   301  TGTTCCGATCACCTGAGAAACCGTCAGGTGGTTGGTCTGTGCCGGATCAGGAAT  360
          |||||||||||||||||||||||||||||||||||||||||||
Db   301  TGTTCCGATCACCTGAGAAACCGTCAGGTGGTTGGTCTGTGCCGGATCAGGAAT  360

Qy   361  TCGGAGAATCCCCGTGGTCGTTGTTTCTTTTCCTGAAATCTGCAAGTCCCAGCAGCAGCAG  420
          |||||||||||||||||||||||||||||||||||||||||||
Db   361  TCGGAGAATCCCCGTGGTCGTTGTTTCTTTTCCTGAAATCTGCAAGTCCCAGCAGCAGCAG  420

Qy   421  AGCAAGAGCAATGGCCGTGCAGGGAGTTTGATCAGTTGATTACTTTGATGCACTAGTACTAGGCG  480
          |||||||||||||||||||||||||||||||||||||||||||
Db   421  AGCAAGAGCAATGGCCGTGCAGGGAGTTTGATCAGTTGATTACTTTGATGCACTAGTACTAGGCG  480

Qy   481  TTCGTTCCATGTCGCTCTCACGCGTGCCATAGTGCCATAGATCCTGCATCAATCATCAGGCC  540
          |||||||||||||||||||||||||||||||||||||||||||
Db   481  TTCGTTCCATGTCGCTCTCACGCGTGCCATAGTGCCATAGATCCTGCATCAATCATCAGGCC  540

Qy   541  AAGATTATATTCCTCACATTTTTCTTCCTATCGCTCAGTCGTCGTCGTTGGGAGCTTA  600
          |||||||||||||||||||||||||||||||||||||||||||
Db   541  AAGATTATATTCCTCACATTTTTCTTCCTATCGCTCAGTCGTCGTCGTTGGGAGCTTA  600

Qy   601  AAAATTATGAAAAAGCAGCTGCTAGCTGGTGAGAAGCTAGCTAGGAGAATTGAGTTCT  660
          |||||||||||||||||||||||||||||||||||||||||||
Db   601  AAAATTATGAAAAAGCAGCTGCTAGCTGGTGAGAAGCTAGCTAGGAGAATTGAGTTCT  660

Qy   661  ACGTTCATTCTCCAGATTCTACAAATTACACAGATTCTTATAATTTAGGTAAAAAGCTGGACT  720
          |||||||||||||||||||||||||||||||||||||||||||
Db   661  ACGTTCATTCTCCAGATTCTACAAATTACACAGATTCTTATAATTTAGGTAAAAAGCTGGACT  720

Qy   721  GTTTGGGAGCTTCTGTCACTCTGATCGATTCGTGAGAAGCTGCTAGAAGCTTCCC  780
          |||||||||||||||||||||||||||||||||||||||||||
Db   721  GTTTGGGAGCTTCTGTCACTCTGATCGATTCGTGAGAAGCTGCTAGAAGCTTCCC  780

Qy   781  CAAACAGACACCCCTAGTTGTTGTACTCTAGTCGGCAATTCACTCTATTTTATATAACCTTGC  840
          |||||||||||||||||||||||||||||||||||||||||||
Db   781  CAAACAGACACCCCTAGTTGTTGTACTCTAGTCGGCAATTCACTCTATTTTATATAACCTTGC  840

Qy   841  TCTCTAGCTTATCAAACGTAGCCAAGACTTGAATTTAAAGCTTAAAATTGATTTGATGTGT  900
          |||||||||||||||||||||||||||||||||||||||||||
Db   841  TCTCTAGCTTATCAAACGTAGCCAAGACTTGAATTTAAAGCTTAAAATTGATTTGATGTGT  900

Qy   901  TCTTTTCATCGTAATTCACTTACCACCGGCATTGAATTTAAAAATAATTT  960
          |||||||||||||||||||||||||||||||||||||||||||
Db   901  TCTTTTCATCGTAATTCACTTACCACCGGCATTGAATTTAAAAATAATTT  960

Qy   961  TTAGAGCTGATTTGATTTTTCAGCGGCGAGTAAGCATTAGTGTTATTTTCACGTATGTAAATCTGG  1020
          |||||||||||||||||||||||||||||||||||||||||||
Db   961  TTAGAGCTGATTTGATTTTTCAGCGGCGAGTAAGCATTAGTGTTATTTTCACGTATGTAAATCTGG  1020

Qy  1021  CCTATAAATTATTAATTTAATTTCAGCGGAGTAAGCATTAGTGTTATGGTTATAATGGTTATAATT  1080
          |||||||||||||||||||||||||||||||||||||||||||
Db  1021  CCTATAAATTATTAATTTAATTTCAGCGGAGTAAGCATTAGTGTTATGGTTATAATGGTTATAATT  1080

Qy  1081  TATGCTTAAATCTCTTTACTTGGACTTAGTGGGACAAATTCGTTAAATGCAATTCTTCGTGCCA  1140
          |||||||||||||||||||||||||||||||||||||||||||
Db  1081  TATGCTTAAATCTCTTTACTTGGACTTAGTGGGACAAATTCGTTAAATGCAATTCTTCGTGCCA  1140

Qy  1141  TCTCTATAATACGGGCCGTAGCTTTGCTCTTGTATATCGTACACAAGAGAACTAGGCTGGCAA  1200
          |||||||||||||||||||||||||||||||||||||||||||
Db  1141  TCTCTATAATACGGGCCGTAGCTTTGCTCTTGTATATCGTACACAAGAGAACTAGGCTGGCAA  1200

Qy  1201  AGTCCTCAAGGCGGAACCGCGCCTCCATCTTCTCCTTCCAGCTGTCTCTGTATCTGCACAAGAACTAGCTGGCAA  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  AGTCCTCAAGGCGGAACCGCGCCTCCATCTTCTCCTTCCAGCTGTCTCTGTATCTGCACAAGAACTAGCTGGCAA  1260

Qy  1261  CGCCATCGGCCATCGCGCCATGGCTCTCATGGTGCCAGAGATATCCAGACGATCTGAAATTAACT  1320
          |||||||||||||||||||||||||||||||||||||||||||
Db  1261  CGCCATCGGCCATCGCGCCATGGCTCTCATGGTGCCAGAGATATCCAGACGATCTGAAATTAACT  1320

Qy  1321  TGACGCATATCATCTCATCATCATCATCAGTTCAATTCTCTGGATTGTGCAGAACATGTTTA  1380
          |||||||||||||||||||||||||||||||||||||||||||
Db  1321  TGACGCATATCATCTCATCATCATCATCAGTTCAATTCTCTGGATTGTGCAGAACATGTTTA  1380

Qy  1381  GCTGATCAGTAAACGTTGCCGTGTTGAAATTGGTCCTTTATCAGGTGGTTCAACCTGG  1436
          |||||||||||||||||||||||||||||||||||||||||||
Db  1381  GCTGATCAGTAAACGTTGCCGTGTTGAAATTGGTCCTTTATCAGGTGGTTCAACCTGG  1436
```

RESULT 3

LOCUS   AP005392   145828 bp   DNA   linear   PLN 28-JUL-2004
DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 9,
            PAC clone:P0463D04.
ACCESSION   AP005392
VERSION   AP005392.3 GI:50725855
KEYWORDS   .
SOURCE   Oryza sativa (japonica cultivar-group)
  ORGANISM Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   Sasaki,T., Matsumoto,T. and Katayose,Y.
  TITLE   Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, PAC
            clone:P0463D04
  JOURNAL   Published Only in Database (2002)
REFERENCE   2 (bases 1 to 145828)
  AUTHORS   Sasaki,T., Matsumoto,T. and Katayose,Y.
  TITLE   Direct Submission
  JOURNAL   Submitted (06-JUN-2002) Takuji Sasaki, National Institute of
            Agrobiological Sciences, Rice Genome Research Program; Kannondai
            2-1-2, Tsukuba, Ibaraki 305-8602, Japan
            (E-mail:tsasaki@nias.affrc.go.jp,
            Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT   On Jul 27, 2004 this sequence version replaced gi:49169760.
            Genes were predicted from the integrated results of the following:
            GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH
            (http://www.softberry.com/), GeneMark.hmm
            (http://opal.biology.gatech.edu/GeneMark/), GlimmerM
            (http://www.tigr.org/tdb/glimmerm/glnr_form.html), RiceHMM
            (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor
            (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4
            (http://globin.cse.psu.edu/html/docs/sim4.html), gap2
            (http://www.tigr.org/software/glimmerm/), BLASTN and BLASTX. The
            genomic sequence was searched against NCBI NonRedundant Protein
            database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
            sequence database at RGP or DDBJ. Protein homologies of the coding
            regions were searched against NCBI NonRedundant Protein database
            with BLASTP. ESTs represent the identified cDNA sequences using
            BLASTN with the corresponding DDBJ accession no. and RGP clone ID.
            Full-length cDNAs represent the identified cDNA sequences using
            BLASTN with the corresponding DDBJ accession no.
            A gene with identity or significant homology to a protein is
            classified based on the protein name to indicate the homology level
            such as same name, 'putative-' and '-like protein'. A gene without
            significant homology to any protein but with full-length cDNA or
            EST homology (covering almost the entire length of partial
            sequence) is classified as an 'unknown' protein. A gene predicted
            by two or more gene prediction programs is classified as a
            'hypothetical' protein according to IRGSP standard. A gene

```
           predicted by a single gene prediction program is also classified as
           a probable 'hypothetical' protein and is included as a
           miscellaneous feature of the sequence.

FEATURES             Location/Qualifiers
     source          1..145828
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="9"
                     /clone="P0463D04"
                     /note="This sequence of P0463D04 is from SP6 to T7 of the PAC clone.
                     The orientation of the sequence is from SP6 to T7 of the PAC clone.
                     This sequence of P0463D04 has an overlap with P0463G11(DDBJ:
                     AP005633) clone at 5' end and with
                     OSJNBa0087J09(DDBJ: AC108761) clone at 3' end. Detailed information
                     on overlap and assembly quality together with annotation of this
                     entry is available at
                     http://rgp.dna.affrc.go.jp/GenomeSeq.html."
     gene            join(5377..5591,5706,6535..6697)
                     /gene="P0463D04.1"
     mRNA            join(5377..5591,5706,6535..6697)
                     /gene="P0463D04.1"
     misc_feature    join(5377..5591,5706,6535..6697)
                     /gene="P0463D04.1"
                     /note="hypothetical ORF
                     predicted by RiceHMM
                     this category is not included in IRGSP standard"
     gene            complement(9919..13083)
                     /gene="P0463D04.2"
     mRNA            complement(join(<9919..10010,10314..>13083))
                     /gene="P0463D04.2"
                     /note="start and end point are not identified"
     CDS             complement(join(9919..10010,10314..13083))
                     /gene="P0463D04.2"
                     /note="contains full-length cDNA(s): AK064218"
                     /codon_start=1
                     /product="putative PPR1"
                     /protein_id="BAD33385.1"
                     /db_xref="GI:50725856"
                     /translation="MAEIAVLFVIKIGIAVAGDTLKLAIPLFAKKYELKKVELVTAL
                     PVNMRQIKKELEIINAFLKELGMNGYKGEVVETWIRQVRRLAHDMEDVVDEMYVVGK
                     NKHKKSMAVKKIKKKPKPLFSLDEIATKADMINTELVELSKRLDRWTRPLSSGIIYP
                     PTNYNSEQQLYLPGYDYSINDNELVGIDKNRQTLIESLRLIEDCSLRIIAVWGWGGLGK
                     STLVNDIYKNEAIVSNNCHAWLCISQSSKMHDIWQNMLKELCGEDNRGVDAENMNNR
                     ELRLELAKILRQRYLIILDDVWLAADLLKIRKVLVDNGLGSRVIITRIREVASIAB
                     DGCKIRLEPLNNHDAWLLFCRKAFPKIENHMCPPELHQCGMDIVNKCGGLPLALVTIG
                     SILSLKPRNKKGWRLFYNQLISEVHNNENLNRVEKIINLSYKHLPNYLKNCFLYCAMF
                     PEDYIIQRKRLIRLWIAEGFIEQKGTGSLEDVABGFIYTELVARSMIQVVARNSFRIQ
                     CLRMHDILRELAIFQSKKEESFSTYYDDHGVVOVGSDSRRVSVLQCNSEIRSTVDPSR
                     LRTFLAFDTSMALSSASYFIFSESKYLAVLELSGLPIETIPYSVGELFNLRYLCLNDT
                     NVKEFPKSITKLLNLQTLSLERTQOLLNFPRGFSNLKKLRHLLVWKLVDATYKSLNNWE
                     SLEPFBGIWNLKELQSLCEVRATRDFVSKLGNLSQLRSLCITYVRSSHICAQLCNSLSK
                     MQHLTRLHIRANWEDEVLLLDDLMLPNPLRKLDLLGQLSKGTLESPFFTHGNELLQL
                     ELSRCQLTVNLVAWLSKUSNLTELRLTRVYTGQOLSFHANCFPNLKKALLWDLQOVNQ
                     IYIQEGALSSLQYLHIDSLMELRDVPTGIFLRSVKRAYFTMHSDFVRNLRTGKVNH
                     IPKVYWSTQGVSAEPANLPGESSTNPQWRMLGGSGWVPI"
     gene            complement(14811..15236)
                     /gene="P0463D04.3"
     mRNA            complement(join(14811..14903,14991..>15236))
                     /gene="P0463D04.3"
                     /note="start and end point are not identified"
     CDS             complement(join(14811..14903,14991..15236))
                     /gene="P0463D04.3"
                     /note="hypothetical protein"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAD33386.1"
                     /db_xref="GI:50725857"
                     /note="predicted by GeneMark.hmm etc."
                     /translation="MTGDHPGFRRGCSSRQOPPPPPAGHLHAAGALCSLTRPYAACA
                     RQGGLYRRHRHRRGCPRRKPLPPPAAFAITPPRSTEHDRIHGEARFISVEDGEIFG
                     OLGVLTLAKS"
     gene            18801..20386
                     /gene="P0463D04.4"
     mRNA            18801..20386
                     /gene="P0463D04.4"

     CDS             18838..20235
                     /note="supported by full-length cDNA(s): AK109585"
                     /gene="P0463D04.4"
                     /note="contains full-length cDNA(s): AK109585"
                     /codon_start=1
                     /product="phosphofructokinase-like"
                     /protein_id="BAD33387.1"
                     /db_xref="GI:50725858"
                     /translation="MDDGCETPMPSLSHLHKLPTLAGPAAIPNPIARHPLYHPSPSFFIS
                     PFDVVLRDILFDASPASAGEPRRRVAYHRAGPRRRVAFDPATVRAAIFTCGGLCP
                     GINTVVRELVVGLSELYCGVRGGVFGVGVRNGYRGPYSDEVVPLDDAAVEHWHKAGGAALG
                     TSRGGFDLARIVDAIEREGFNQVTAVGGDGTMRGAARIHREVRRGRLLAVAVAGIPKT
                     VDNDVGVVDRSFGFHTAVEAAOOAIAAGHVEAESAANGVGLVKLMGRSAGHIALHATL
                     SSRDVDCCLIPBEDFYLRGAGGLFDFLYRRIKDNGHAVVVAEGAGQRLIPRTTTSA
                     SGACAGADESGNETFLDVGAWLKAEMRAWWEEEHAGEVFTVKVIDPTYMIRAVPANAG
                     DNIYCTLLAHAAIHGAMAGYTGFVSGTINGNYAYIPMDEVARARXNPVDTKDHKMAWVR
                     SIINQPDFIRAGPTS"
     gene            complement(22101..29653)
                     /gene="P0463D04.5"
     misc_feature    complement(22101..29653)
                     /gene="P0463D04.5"
                     /note="probably inactive due to including frameshift(s) in
                     CDS
                     probably inactive due to including stop codon(s) in CDS
                     pseudogene, subtilisin-like protein"
                     /pseudo
     CDS             complement(join(30660..30826,30990..31080))
                     /gene="P0463D04.6"
                     /note="contains full-length cDNA(s): AK062866"
                     /codon_start=1
                     /product="unknown protein"
                     /protein_id="BAD33388.1"
                     /db_xref="GI:50725859"
                     /translation="MCCPSKACCICTLIVLIVAVGIVFGGVYTRGFHKLTSSIHD
                     ASSTAAGSLRAHOFFAPPPY"
     gene            complement(30660..31080)
                     /gene="P0463D04.6"
     mRNA            complement(join(30660..30826,30990..31080))
                     /gene="P0463D04.6"
                     /note="contains full-length cDNA(s): AK062866"
     misc_feature    complement(31283..31731)
                     /gene="P0463D04.7"
                     /note="hypothetical ORF
                     predicted by GENSCAN
                     this category is not included in IRGSP standard"
     gene            complement(31283..31731)
                     /gene="P0463D04.7"
     mRNA            complement(31283..31731)
                     /gene="P0463D04.7"
     gene            32587..32862
                     /gene="P0463D04.8"
     misc_feature    32587..32862
                     /gene="P0463D04.8"
                     /note="hypothetical ORF
                     predicted by GlimmerM
                     this category is not included in IRGSP standard"
     gene            34608..35096
                     /gene="P0463D04.9"
     mRNA            <34608..>35096
                     /gene="P0463D04.9"
                     /note="start and end point are not identified"
     CDS             34608..35096
                     /gene="P0463D04.9"
                     /note="contains EST(s): AU077507(C52306),AU077506(C52306)
                     similar to Oryza sativa chromosome 8, B1142B04.28"
                     /codon_start=1
                     /product="unknown protein"
                     /protein_id="BAD33389.1"
                     /db_xref="GI:50725860"
                     /translation="MDRRRRGGQPERKAAAATTKKRGRDEEVAAAAPPPCGAVDFPFE
                     BGDEEAAADVGDQAPGVFRFPFWQSCRGGLGVVAGGGGGGGGWEMRDVFFRSLVDGG
                     AAAIGVPGDRLVSPPPAELRRALFDDVGAWLAAAAGDGEVDPVWRSALLIEGEPPAPA
                     TV"
     gene            36525..41570
                     /gene="P0463D04.10"
```

misc_feature   36525..41570
    /gene="P0463D04.10"
    /note="probably inactive due to including stop codon(s) in
    CDS

Query Match   90.7%;   Score 1303;   DB 8;   Length 145828;
Best Local Similarity   97.5%;   Pred. No. 6.8e-294;
Matches 1409;   Conservative   0;   Mismatches   25;   Indels   11;   Gaps   8;

Qy   1 CATTCAGAAATCATCTCCAGCCTACAATGTACTCTCTCCCCAAATACAAGTGTCTCTATGA 60
             |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   62886 CATTCAGAAATCATCTCCAGCCTACAATGTACTCTCTCTCCCCAAATACAAGTGTCTCTATGA 62945

Qy   61 TTCAAAATTTGTCCTACAAATATAAAACATTTCCAGCATGAAATCCATACACACAAGTGTC 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62946 TTCAAAATTTGTCCTACAAATATAAAACATTTCCAGCATGAAATCCATACACACAAGTGTC 63005

Qy   121 CTAATCAGATGCTTGGAGGGAAAAAATCTAAGCGGATTCCAATATGCAAAAAATTGATCACTGA 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63006 CTAATCAGATGCTTGGAGGGAAAAAATCTAAGCGGATTCCAATATGCAAAAAATTGATCACTGA 63065

Qy   181 AGTAACTGAAAAGAGAATATCTCGTTTTAACATTAGTGTTATTTTATTTGAACACACTAAA 240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63066 AGTAACTGAAAAGAGAATATCTCGTTTTAACATTAGTGTTATTTTATTTGAACACACTAAA 63125

Qy   241 AAAATTGTTTATATATTTTAGTACAAAATCGAGTAGTAGGTAAGGTAAGGATCG 300
             ||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   63126 AAAATTGTTTATATATTTTAGTACAAAATCGAGTAGTAGGTAGTAGGATCG 63185

Qy   301 TGTTCCGATCACCTGAGAAACCGTCAGGTGGTTTGTCTGTCCGTCCGATCAGAAT 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63186 TGTTCCGATCACCTGAGAAACCGTCAGGTGGTTTGTCTGTCCGTCCGATCAGAAT 63245

Qy   361 TCGGAGATCCGCCGTCGTTTCTTTCTTGAAATCTGCAAGTCCCAGCAGCAGCAGCAG 420
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63246 TCGGAGATCCGCCGTCGTTTCTTTCTTGAAATCTGCAAGTCCCAGCAGCAGCAGCAG 63305

Qy   421 AGCAAGAGCAAATGGCCGTGCAGGGAGTTTGATGATCGTAGCTAGCTACTAGGCG 480
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63306 AGCAAGAGCAAATGGCCGTGCAGGGAGTTTGATGATCGTAGCTAGCTACTAGGCG 63365

Qy   481 TTCGTTCCATGTCGCTCTCACGCCGTGCGAATGTGCCATGACATCCATCATCGCC 540
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63366 TTCGTTCCATGTCGCTCTCACGCCGTGCGAATGTGCCATGACATCCATCATCGCC 63425

Qy   541 AAGATTATTAATTCCTCCACATTTTTCTTCCTTCTCGCTCGTCGTTTTGGGAGCTTA 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63426 AAGATTATTAATTCCTCACATTTTTCTTCCTTCTCGCTCGTCGTTTTGGGAGCTTA 63485

Qy   601 AAAATTATGAAAAAGCAGCTGCTGGTAGGCTAGCTGGTGAGAATCTGAAGAATTGAGTTCT 660
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63486 AAAATTATGAAAAAGCAGCTGCTGGTAGGCTAGCTGGTGAGAATCTGAAGAATTGAGTTCT 63545

Qy   661 ACGTTCATTCTCCAGATTCTACAATTACAGATTCTTATATTTAGTAAAAAAGCTCGGACT 720
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63546 A-GTTCATTCTCCAGATTCTACAATTACAGATTCTTATATTTAGTAAAAAAGCTCGGACT 63604

Qy   721 GTTTGGGGAGCTTCTGTCAGCGCGGAGATTCTGTGAGAAGTCGCAGCTGCGAAGCTTCCC 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63605 GTTTGGGGAGCTTCTGTCAGCGCGGAGATTCTGTGAGAAGTCGCAGCTGCGAAGCTTCCC 63664

Qy   781 CAAACAGACCCCTAGTTGTACTCTAGCTGATCGATTCACTCTATTTTATATACACCTTGC 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63665 CAAACAGACCCCTAGTTGTACTCTAGCTGATCGATTCACTCTATTTTATATGCACCTTGC 63724

Qy   841 TCTCTAGCTTATCAAACGTAGCCAAGACTTGAATTTAAAAGCTTAAAATTGATTTTGATGT 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63725 TCTCTAGCTTATCAAACGTAGCCAAGACTTGAATTTAAAAGCTTAAAATTGATTTTGATGT 63784

Qy   901 TCTTTTTCATCGTTAATTCACTTACCGACCTTAGTCGGCACCATTTGAATTTTTTAAAAATAATTT 960
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63785 TCTTTTTCATCGTTAATTCACTTACCGACCTTAGTCGGCACCATTTGAATTTTTTAAAAATAATTT 63844

Qy   961 TTAGAGCTGATTTTTGATTTTTTTTTTCAGCGGAGTAAGCGGAAATTTATTTTCACGTAGTGTAAAAGTTTA 1020
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63845 TTAGAGCTGATTTTTGATTTTTTTTTTCAGCGGAGTAAGCGGAAATTTATTTTCACGTAGTGTAAAAGTTTA 63904

Qy   1021 CCTATAAATTATTAATTTCACGCGAGTAGCATTAAGTGGGAGGTTATGGGTTATAATCATCTGG 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63905 CCTATAAATTATTAATTTCACGCGAGTAGCATTAAGTGGGAGGTTATGGGTTATAATCATCTGG 63964

Qy   1081 TATGCTT-AAATCTCTTTACTTGGACTTAGTTGGGACTTAGTTGGGTAAATGCATTCTCGTTG-C 1138
             ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63965 TATGCTTAAAAATCTCTTTACTTGGACTTAGTTGGGACTTAGTTGGGTAAATGCATTCTCGTTGCC 64024

Qy   1139 CATCTCTTAT-AATACGGCCGCTAGCTTTGCTCTTGTATCTCGCACACAAGAACTAGCTG 1197
             ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   64025 CATCTCTTATAAATACGGCCGCTAGCTTTGCTCTTGTATCTCGCACACAAGAACTAGCTG 64083

Qy   1198 CAAAGTCCTCAAGGCGGAACCGCCCATCTCTCCTTCCAGTCGTTGCACTGGTTCCTT 1257
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   64084 CAAAGTCCTCAAGGCGGAACCGCCCATCTCTCCTTCCAGTCGTTGCACTGGTTCCTT 64143

Qy   1258 CGTCGCCATCGCCATCCCATGGCTCTCATCAGTGTGCAGAGAGATATCC-----AGCATGCTGAA 1313
             ||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   64144 CGTCGCCATCGCCATCCCATGGCTCTCATCAGTGTGCAGAGAGATATCCAGCATGCTGAA 64203

Qy   1314 ATTAACTTGACGGCATATCANTCTCATCACTTGCATTCAACTTCTTGGGATTGTGCAGAC 1373
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   64204 ATTAACTTGACGGCATATCACTCTCATCACTTGCATTCAACTTCTTGGGATTGTGCAGAC 64263

Qy   1374 ATGTTTAGCTGAT-CAGTAAAACGTTGCCGTGTTGAATTGGTCCCTT-ATCAGGTGGTTCAA 1431
             |||||||||||| ||||||||||||||||||||||||||||||||| ||||||||||
Db   64264 ATGTTTAGCTGATGCAGTAAAACGTTGCCGTGTTGAATTGGTGTCCCTTGATCAGGTGGTTCAA 64323

Qy   1432 CCTGG 1436
             |||||
Db   64324 CCTGG 64328

RESULT 4
AP005633
LOCUS   AP005633   154188 bp   DNA   linear   PLN 15-SEP-2004
DEFINITION   Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 9,
               PAC clone:P0463G11.
ACCESSION   AP005633
VERSION   AP005633.3   GI:50726137
KEYWORDS   .
SOURCE   Oryza sativa (japonica cultivar-group)
  ORGANISM   Oryza sativa (japonica cultivar-group)
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
               Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   Sasaki,T., Matsumoto,T. and Katayose,Y.
  TITLE   Oryza sativa nipponbare(GA) genomic DNA, chromosome 9, PAC
               clone:P0463G11.
  JOURNAL   Published Only in Database (2002)
REFERENCE   2   (bases 1 to 154188)
  AUTHORS   Sasaki,T., Matsumoto,T. and Katayose,Y.
  TITLE   Direct Submission
  JOURNAL   Submitted (07-AUG-2002) Takuji Sasaki, National Institute of
               Agrobiological Sciences, Rice Genome Research Program; Kannondai
               2-1-2, Tsukuba, Ibaraki 305-8602, Japan
               (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
               Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT   On Jul 27, 2004 this sequence version replaced gi:41152739.
               Genes were predicted from the integrated results of the following:
               GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH
               (http://www.softberry.com/), GeneMark.hmm
               (http://opal.biology.gatech.edu/GeneMark/), GlimmerM
               (http://www.tigr.org/tdb/glimmerm/glmr_form.html), RiceHMM
               (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor
               (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4
               (http://globin.cse.psu.edu/html/docs/sim4.html), gap2
               (http://www.tigr.org/software/glimmerm/), BLASTN and BLASTX. The

genomic sequence was searched against NCBI NonRedundant Protein
database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
sequence database at RGP or DDBJ. Protein homologies of the coding
regions were searched against NCBI NonRedundant Protein database
with BLASTP. ESTs represent the identified cDNA sequences using
BLASTN with the corresponding DDBJ accession no. and RGP clone ID.
Full-length cDNAs represent the identified cDNA sequences using
BLASTN with the corresponding DDBJ accession no.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-', and '-like protein'. A gene without
significant homology to any protein but with full-length cDNA or
EST homology (covering almost the entire length of partial
sequence) is classified as an 'unknown' protein. A gene predicted
by two or more gene prediction programs is classified as a
'hypothetical' protein according to IRGSP standard. A gene
predicted by a single gene prediction program is also classified as
a probable 'hypothetical' protein and is included as a
miscellaneous feature of the sequence.
The orientation of the sequence is from SP6 to T7 of the PAC clone.
This sequence of P0463G11 has an overlap with P0556A05(DDBJ:
AP005759) clone at 5' end and with P0463D04(DDBJ: AP005392) clone
at 3' end. Detailed information on overlap and assembly quality
together with annotation of this entry is available at
http://rgp.dna.affrc.go.jp/GenomeSeq.html.

FEATURES   Location/Qualifiers
source   1..154188
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="9"
                     /clone="P0463G11"
gene   1..3601
                     /gene="P0463G11.1"
mRNA   join(1..52,1033...2015,2554..3601)
                     /gene="P0463G11.1"
                     /note="supported by full-length cDNA(s): AK100536"
CDS   join(1144..2015,2554..3433)
                     /gene="P0463G11.1"
                     /note="contains EST(s): D15569(C0846A),C22471(C0846)
                     contains full-length cDNA(s): AK100536"
                     /codon_start=1
                     /product="putative Protein kinase G11A"
                     /protein_id="BAD46322.1"
                     /db_xref="GI:52077280"
                     /translation="MSSKSPLQSMPGQAAKPKGTGEQGQGDAEVVSERTPSDQHCGSP
                     ASTLDKELSGLSSESGILDDSLVAEEDSGELKEIQNLDCNGNQEKKTSQKSSTSDSF
                     ASAKVSDGTNSLRKTSGSAKTSDRANFTESGKSSMCRASTSDSDSRSSRSMSSATT
                     KPHKGNDSRWEAIQTVKTKDGILGLNHFRLLKKLGSGDIGSVYLSELSGTRSYFAMKV
                     MDKASLASRKKLLRSQTELRLQSILDHPFLPTLYTHFRTDKFSCLVMEFCPGGDLHTL
                     RQRQPGKHFSEQAAKFYVAEVLLALEYLHMLGIIYRDLKPENVLVREDGHIMLSDFDL
                     SLRCSVSPTVIKSANPGIDALQRNNAAYCVQPACIEPSCIQPSCVAPTTCRGPRFFKS
                     KSKSKSKKKSKPEAANQASLFPELIAEPTDARSMSFVGTHEYLAPEIIKGEGHGSAV
                     DWWTFGIFLYELLFGKTPFKGSGNRATLFNVGQPLRPEPSPIVSFSARDMIRGLLVK
                     DPQHRLGYKRGATEIKQHPFFEGVNWALIRCASPPDIPKPVELDCREKQVPSTNGKVA
                     PASNQKGSDNYLPFFF"
gene   complement(6639..7305)
                     /gene="P0463G11.2"
mRNA   complement(join(6639...6667,7140..>7305))
                     /gene="P0463G11.2"
                     /note="start and end point are not identified"
CDS   complement(join(6639...6667,7140..>7305))
                     /gene="P0463G11.2"
                     /note="predicted by RiceHMM etc."
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAD46323.1"
                     /db_xref="GI:52077281"
                     /translation="MSNNKGGAPVHTPTHVLAANGRCQTKFASLLRRRIQVDIYDLSS
                     VSVGRRSVGLRSSEETLLSVV"
gene   8141..9067
                     /gene="P0463G11.3"

misc_feature   8141..9067
                     /gene="P0463G11.3"
                     /note="probably inactive due to including stop codon(s) in
                     CDS"
                     /pseudo
                     /note="pseudogene, putative cyst nematode resistance protein"
gene   complement(11803..12411)
                     /gene="P0463G11.4"
mRNA   complement(<11803...>12411)
                     /gene="P0463G11.4"
                     /note="supported by full-length cDNA(s): AK109529"
misc_feature   complement(11803..12411)
                     /gene="P0463G11.4"
                     /note="contains full-length cDNA(s): AK109529
                     non-coding transcript
                     probably inactive due to 5'exon missing in CDS"
gene   join(13029..13245,14127..14276)
                     /gene="P0463G11.5"
mRNA   join(13029..13245,14127..14276)
                     /gene="P0463G11.5"
misc_feature   join(14181,14255..14276)
                     /gene="P0463G11.5"
                     /note="hypothetical ORF
                     predicted by GENSCAN
                     this category is not included in IRGSP standard"
gene   16195..17553
                     /gene="P0463G11.6"
mRNA   join(<16195...16722,17461..>17553)
                     /gene="P0463G11.6"
                     /note="start and end point are not identified"
CDS   join(16195..16722,17461..17553)
                     /gene="P0463G11.6"
                     /note="predicted by GENSCAN etc."
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAD46324.1"
                     /db_xref="GI:52077282"
                     /translation="MEGHHGEGNVCEREGEDGEBMGPQMGAWLVLVGYVGRPGLQCWGG
                     GGTWPRAHEAKPRPLTCAASGVAFRRRPCGAVHPARQVRWAPPPROVTWSRGGSSRA
                     GERRRRGGGGLVLCPGAVNLANHQEBAAIAVAVARQCCAAAGSAASRRWNRIGRRF
                     AAGSGGRGASESRAAPGDGSDVCACVSTHHQRASEVDGDEDGDYD"
gene   complement(19698..20346)
                     /gene="P0463G11.7"
mRNA   complement(<19698...>20346)
                     /gene="P0463G11.7"
                     /note="supported by full-length cDNA(s): AK107719"
misc_feature   complement(19698..20346)
                     /gene="P0463G11.7"
                     /note="contains full-length cDNA(s): AK107719
                     non-coding transcript
                     probably inactive due to 5'exon missing in CDS"
gene   20526..20981
                     /gene="P0463G11.8"
mRNA   <20526...>20981
                     /gene="P0463G11.8"
                     /note="start and end point are not identified"
CDS   <20526...>20981
                     /gene="P0463G11.8"
                     /note="predicted by GlimmerM etc."
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAD46325.1"
                     /db_xref="GI:52077283"
                     /translation="MLLPRDADLPRLPTSTRRLFRSPLPVPEAYSIKPVSYQGKRNQY
                     SARAEETKPEAKAYFWRREEATTGPGGPLAGPWDAEAHTTRRRGGGGEELPPTNGGG
                     CARSSSDRAPVSLPPILSPRLCSRGALIACLILGWFGCFVSRSVCNGG"
gene   21597..22709
                     /gene="P0463G11.9"
mRNA   join(<21597...22565,22620..>22709)
                     /gene="P0463G11.9"
                     /note="start and end point are not identified"
CDS   join(21597..22565,22620..22709)
                     /gene="P0463G11.9"
                     /note="predicted by RiceHMM etc."
                     /codon_start=1

```
                /product="hypothetical protein"
                /protein_id="BAD46326.1"
                /db_xref="GI:52077284"
                /translation="MLATRCALLFPSRSVSSHHPSDSVLIDEAFSQSTLAVVIGPD
                DFLCGSSNIGLVCLYTQTTTIKIINLATGGCHLHEKPAKNLAGDYSSLYSFGFRPVTK
                KYKVTHFLGDSRKAHPRAXDSFYAIQVHTLGDEKWKDVGSPEALSLNCVKNSGVVNVD
                GIMYWLTEDQGASWQHAVISFDLSRESFGRIQLPTVVLEDFAPYYGPHQYWIKEIDGKV
                CIATSQTTQNQPRELIGEIQIWTLHIHLEQRWIQKMIQSSPQCILGPNIFHGDKILS
                QQYGSILYSCELLGKNLEVKMSNADRLLDFTPRKPGNMQSYTPVKSLSGWELKKWEAW
                ESNRCVLEDTWKDVQHLE"

gene        24677..28278
                /gene="P0463G11.10"
mRNA        join(24677..24760,25402..26693,26786..26845,26968..27060,
                27411..27467,28057..28278)
                /gene="P0463G11.10"
CDS         join(25431..26693,26786..26845,26968..27060,27411..27467,
                28057..28095)
                /gene="P0463G11.10"
                /note="contains full-length cDNA(s): AK063439"

Query Match   90.7%;  Score 1303;  DB 8;  Length 154188;
Best Local Similarity  97.5%;  Pred. No. 6.8e-294;
Matches 1409;  Conservative  0;  Mismatches 25;  Indels 11;  Gaps  8;

Qy        1 CATTCAGAATCATCTCCAGCCTACAAGTACGTCTCTCCCAATAATACAAGTGTCTCTATGA 60
            ||||||||||||||||||||||||||| ||||| ||||||||||||||||||||||||||||
Db   115748 CATTCAGAATCATCTCCAGCCTACAATGTACTCTCTCCCAATAATACAAGTGTCTCTATGA 115807

Qy       61 TTCAAAATTGTCCTACAATATAAACATTCCAGCAATGTACTCTCTCCCAATAATTTTTCAG 120
            |||||||||||||||||| ||||||||||||||||||||||| | ||||||||||||||
Db   115808 TTCAAAATTGTCCTACAAATAAAACATTCCAGCAATGTACATACATTAATTTTTTCAG 115867

Qy      121 CTAATCAGATGCTTGGAGGGGAAAAATCTAAGGCGATTCAATATGCAAAAAATTGATCACTGA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   115868 CTAATCAGATGCTTGGAGGGGAAAAATCTAAGGCGATTCAATATGCAAAAATTGATCACTGA 115927

Qy      181 AGTAACTGAAAGAGAGAATATCTCGTTTTTAACATTAGTAGTCTAGTATTTTATTTAAACAAACTAAA 240
            ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db   115928 AGTAACTGAAAGAGAGAATATCTCGTTTTTAACATTAGTAGTCTAGTATTTTATTTAAACAAACTAAA 115987

Qy      241 AAAATTGTTTATATTTTAGTACAAAATCGAGTAGCAGAGGCTAGCCGTAAGATCG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   115988 AAAATTGTTTATATTTTAGTACAAAATCGAGTAGCAGAGGCTAGCCGTAAGATCG 116047

Qy      301 TGTTCCGATCACCTGAGAAAACCGTCAGGTGGTTGTCTGTGCCGGATCAGAAT 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   116048 TGTTCCGATCACCTGAGAAAACCGTCAGGTGGTTGTCTGTGCCGGATCAGAAT 116107

Qy      361 TCGGAGGATCCGCCGTCGTTTCTTTCCTGAAATCTGCAAGTCCCAGCAGCGCAGCAG 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   116108 TCGGAGGATCCGCCGTCGTTTCTTTCCTGAAATCTGCAAGTCCCAGCAGCGCAGCAG 116167

Qy      421 AGCAAGAGCAAATGGCGCTGCAGGGAGTTGATACTTTGATCGATCAGCTACTACTAGGGCG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   116168 AGCAAGAGCAAATGGGCGTGCAGGGAGTTGATACTTTGATCGATCAGCTACTACTAGGGCG 116227

Qy      481 TTCGTTCCAAGTCGCTCTCACGCCGTGGCGAATGTGCCATGATCGCATCATCGCC 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   116228 TTCGTTCCAAGTCGCTCTCACGCCGTGGCGAATGTGCCATGATCGCATCATCGCC 116287

Qy      541 AAGATTATATTCCTCACATTTTTCTTCCTATCGCTCCTATCGCTCGTTGTTGGGGAGCTTA 600
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   116288 AAGATTATATTCCTCACATTTTTCTTCCTATCGCTCCTAGTCGCTCGTTGTTGGGGAGCTTA 116347

Qy      601 AAATTATGAAAAGCAGCTGCTGAGAAGCTAGCTGGTGAGAATCTGAAGAATTTGAGTTCT 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   116348 AAATTATGAAAAGCAGCTGCTGAGAAGCTAGCTGGTGAGAATCTGAAGAATTTGAGTTCT 116407

Qy      661 ACGTTCATTCTCCAGATTCTACAATTACAGATTCTTATAATTAGGTATAAAAGCTGGACT 720
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   116408 A-GTTCATTCTCCAGATTCTACAATTACAGATTCTTATAATTAGGTATAAAAGCTGGACT 116466

Qy      721 GTTTGGGGAGCTTCTGTCAGCGCGGAGATTCTGTGAGAAGAGCTGCAGCTGCTAGAAGCTTCC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   116467 GTTTGGGGAGCTTCTGTCAGCGCGGAGATTCTGTGAGAAGAGCTGCAGCTGCTAGAAGCTTCC 116526

Qy      781 CAAACAGGACCCCTAGTTGTACTCTAGCTGATCGATCGATCGTGAGAAGCTTCC 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   116527 CAAACAGGACCCCTAGTTGTACTCCAGCTGATCGATCGATCGTCACTCTATTTTTATATACACCTTGC 116586

Qy      841 TCTCTAGCTTATCAAACGTAGCCCAGACTTGAATTTTAAATTGATTTGATGT 900
            ||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   116587 TCTCTAGCTTATCAAACGTAGCCCAGACTTGAATTTTAAATTGATTTGATGT 116646

Qy      901 TCTTTTCATGTCTAATTCACTTACCGACACCTTAGTCGGCCGCATTTGAATTTTTAAAAATAATTT 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   116647 TCTTTTCATGTCTAATTCACTTACCGACACCTTAGTCGGCCGCATTTGAATTTTTAAAAATAATTT 116706

Qy      961 TTAGAGCTGATTTTGATTTTTTTTTTTCAGCGCGGAGAGTAAGCATTAGTGTTATGGGTTTA 1020
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   116707 TTAGAGCTGATTTTGATTTTTTTTTTTCAGCGCGGAGAGTAAGCATTAGTGTTATGGGTTTA 116766

Qy     1021 CCTATAAAATATTATTAATTTTCAGCGCGGAGAGTAAGCATTAGTGTTATATAATCATCTGG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   116767 CCTATAAAATATTATTAATTTCAGCGGAGAGTAAGCATTAGTGTTATATAATCATCTGG 116826

Qy     1081 TATGCTT-AAATCTTTACTTGGACTAGTGTGGGACAATTCGTAATGCATTCTCGTG-C 1138
            |||||||  ||||||||||||||||||||||||||||||||||||||||||||| |
Db   116827 TATGCTAAAAATCTTTACTTGGACTTAGTGTGGGACAATTCGTAATGCATTCTCGTGCC 116886

Qy     1139 CATCTCTAT-AATACGGCCTGCTAGTTTGCTGTTCGTCTGTATCGTGCACACAAGACTAGCTGG 1197
            ||||||||  |||||||||||||||||||||||||||||||||||||||||||||||||
Db   116887 CATCTCTATAAATACGGCCTGCTAGTTTGCTGTTCGTCTGTATCGTGCACACAAGACTAGCT-G 116945

Qy     1198 CAAAGTCCTCAAGGCGAACCGCCCTCATCTTCTCCTTCCAGCATCTCCCCCATGGCCGTGCCT 1257
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   116946 CAAAGTCCTCAAGGCGAACCGCCCTCATCTTCTCCTTCCAGCATCTCCCCCATGGCCGTGCCT 117005

Qy     1258 CGTCGCCCATGGCCATCGCCATGGCCTCTCACGGTCAGCAGAGATATCC----AGCATGCTGAA 1313
            |||||||||||||||||||||||||||||||||||||||||||||    ||||||||||
Db   117006 CGTCGCCCATCGCCATGGCCATGGCCTCTCAGGTCAGCAGGTATATCCCAGACCATGCTGAA 117065

Qy     1314 ATTAACTTGACGACATATCATCTCATCTCATGCTAATCTTTCTGGATTGTGCAGAC 1373
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   117066 ATTAACTTGACGACATATCATCTCATCTCATCACTTGCATTCAACTTCTGGATTGTGCAGAC 117125

Qy     1374 ATGTTTAGCTGAT-CAGTAAAGCTTGCCGGTGTTGAATTGGTCCTT-ATCAGGTGGTTCAA 1431
            |||||||||||||  |||||||||||||||||||||||||||||| ||||||||||||
Db   117126 ATGTTTAGCTGATGCAGTAAAGCTTGCCGGTGTTGAATTGGTCCTTGATCAGGTGGTTCAA 117185

Qy     1432 CCTGG 1436
            |||||
Db   117186 CCTGG 117190
```

RESULT 5
BD242729

```
LOCUS       BD242729                1196 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  DNA containing rice anther-specific gene and transgenic plants
            transformed thereby.
ACCESSION   BD242729
VERSION     BD242729.1  GI:33052499
KEYWORDS    JP 2002528125-A/3.
SOURCE      Oryza sativa
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 1196)
  AUTHORS   An,G., Jeon,J.S., Chung,Y.Y. and Lee,S.
  TITLE     DNA containing rice anther-specific gene and transgenic plants
            transformed thereby
  JOURNAL   Patent: JP 2002528125-A 3 03-SEP-2002;
            SYNGENTA PARTICIPATIONS AG
COMMENT     OS   Oryza sativa (rice)
```

```
PN   JP 20252812S-A/3
PD   03-SEP-2002
PF   02-NOV-1999 JP 2000579761
PR   03-NOV-1998 KR 1998/46973,19-NOV-1998 KR 1998/50126 PI
GYNHEUNG AN,JONG SEONG JEON, YONG YOON CHUNG,SICHUL LEE PC
C12N15/09,A01H5/00,C07K14/415,C12N15/00
CC   DNA containing rice anther-specific gene
CC   and transgenic plants
CC   thereby
CC   transformed
FH   Key          Location/Qualifiers
FT   source       1..1196
FT                /organism="Oryza sativa"
FT                /mol_type="genomic DNA"
FT                /db_xref="taxon:4530"

FEATURES             Location/Qualifiers
     source          1..1196
                     /organism='Oryza sativa (rice)'.

ORIGIN
Query Match    83.3%;  Score 1196;  DB 6;  Length 1196;
Best Local Similarity  100.0%;  Pred. No. 5.3e-269;
Matches 1196;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     1 CATTCAGAATCATCTCCAGCCTACAAAGTACTCTCCCCATAATACAAGTGTCTCTATGA    60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 CATTCAGAATCATCTCCAGCCTACAAAGTACTCTCCCCATAATACAAGTGTCTCTATGA    60

Qy    61 TTCAAAATTTGTTCCTACAATATAAACATTTCCAGCATGAAAATCCATACACATTAATTTTCAG   120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TTCAAAATTTGTTCCTACAATATAAACATTTCCAGCATGAAAATCCATACACATTAATTTTCAG   120

Qy   121 CTAATCAGATGCTTGGAGGAGGAAAAATCTAAGCGATTCAATATGCAAAAAATTGATCACTGA   180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CTAATCAGATGCTTGGAGGAGGAAAAATCTAAGCGATTCAATATGCAAAAAATTGATCACTGA   180

Qy   181 AGTAACTGAAAGAGAAATATCTCGTTTTAACATTAGTGTCTAGTATTTATTATAACAACTAAA   240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AGTAACTGAAAGAGAAATATCTCGTTTTAACATTAGTGTCTAGTATTTATTATAACAACTAAA   240

Qy   241 AAAATTGTTTTATATTTTAGGTACAAAATCGAGTAGTAGGCAGTAGGACAGAGCTAGCGTAAGATCG   300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 AAAATTGTTTTATATTTTAGGTACAAAATCGAGTAGTAGGCAGTAGGACAGAGCTAGCGTAAGATCG   300

Qy   301 TGTTCCGATCACCTGAGAAACCGTCAGGTGGTTTGTTCGTGGCCGTCCAGCCGATCAGAAT   360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TGTTCCGATCACCTGAGAAACCGTCAGGTGGTTTGTTCGTGGCCGTCCAGCCGATCAGAAT   360

Qy   361 TCGGAGATCCGCCGTCGTTCTTTCTTCCTGAAAATCTGCAAGTCCCACGCAGCAGCGCAGCAG   420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TCGGAGATCCGCCGTCGTTCTTTCTTCCTGAAAATCTGCAAGTCCCACGCAGCAGCGCAGCAG   420

Qy   421 AGCAAGAGACAATGGCCGTGCAGGGAGTTTGATACTTTGATCCACTAGCTACTACTAGGCG   480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AGCAAGAGACAATGGCCGTGCAGGGAGTTTGATACTTTGATCCACTAGCTACTACTAGGCG   480

Qy   481 TTCGTTCCAGTCGCTCTCACGCCGTGCGAATGTGCCATGATCCTGCACTCATCATCGCC   540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TTCGTTCCAGTCGCTCTCACGCCGTGCGAATGTGCCATGATCCTGCACTCATCATCGCC   540

Qy   541 AAGATTATATTCCTCACATTTTTCTTCCTATCGCTCGTCGTGTTGTTGGGAGCTTA   600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 AAGATTATATTCCTCACATTTTTCTTCCTATCGCTCGTCGTGTTGTTGGGAGCTTA   600

Qy   601 AAAATTATGAAAAGCAGCTGCTCGTGAGAAGCTAGCTGTGAAGAATTGAGTTCT   660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 AAAATTATGAAAAGCAGCTGCTCGTGAGAAGCTAGCTGTGAAGAATTGAGTTCT   660

Qy   661 ACGTTCATTCTCCAGATTCTACAATTCTTATAATTTAGGTAAAAAAGCTGGACT   720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 ACGTTCATTCTCCAGATTCTACAATTCTTATAATTTAGGTAAAAAAGCTGGACT   720

Qy   721 GTTTGGGAGCTTCTGTCAGCCGGAGATTCTGTGAGAAGCTGCAGCTGCTAGAAGCTTCCC   780
```

```
Db   721 GTTTGGGAGCTTCTGTCAGCCGGAGATTCTGTGAGAAGCTGCAGCTGCTAGAAGCTTCCC   780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   781 CAAACAGACCCCTAGTTGTACTTCTAGCTGATCGATCCAGTCACTCACTTATTTTATATACAACCTTGC   840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CAAACAGACCCCTAGTTGTACTTCTAGCTGATCGATCCAGTCACTCACTTATTTTATATACAACCTTGC   840

Qy   841 TCTCTAGCTTATCAAAGTAGCCAAGACTGAAGACTTGAATGAAACTTAAAATGATTTGATGT   900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TCTCTAGCTTATCAAAGTAGCCAAGACTGAAGACTTGAATGAAACTTAAAATGATTTGATGT   900

Qy   901 TCTTTTCATCGTAATTCACTTACCGGACCCTTAGTCGGCATTGAATTTTTAAAAATAATT   960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCTTTTCATCGTAATTCACTTACCGGACCCTTAGTCGGCATTGAATTTTTAAAAATAATT   960

Qy   961 TTAGAGCTGATTTTGAATTTTTTTTTCAGCGCGGAAGCTTAGTGTTGTTATGGGTTA   1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TTAGAGCTGATTTTGAATTTTTTTTTCAGCGCGGAAGCTTAGTGTTGTTATGGGTTA   1020

Qy  1021 CCTATAAAATTATTAAATTTTCAGCGCGGAGTAAGCATTAGTGTGTTATGGGTTATAATCATCTGG   1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 CCTATAAAATTATTAAATTTTCAGCGCGGAGTAAGCATTAGTGTGTTATGGGTTATAATCATCTGG   1080

Qy  1081 TATGCTAAATCTCTTTACTTGGACTTAGTTGGGACAAATCGTAATGCATTCTCGTTGCCA   1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 TATGCTAAATCTCTTTACTTGGACTTAGTTGGGACAAATCGTAATGCATTCTCGTTGCCA   1140

Qy  1141 TCTCTATAAATACGGGCCTGCGTAGCTTTGCTCTTGTATCTGCACAAGAACTAGCTG   1196
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 TCTCTATAAATACGGGCCTGCGTAGCTTTGCTCTTGTATCTGCACAAGAACTAGCTG   1196

RESULT 6
LOCUS       AF042275                2757 bp    DNA     linear   PLN 23-MAR-1999
DEFINITION  Oryza sativa anther-specific protein gene, complete cds.
ACCESSION   AF042275
VERSION     AF042275.1  GI:4091009
KEYWORDS    .
SOURCE      Oryza sativa
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 2757)
  AUTHORS   Jeon,J.S., Chung,Y.Y., Lee,S., Yi,G.H., Oh,B.G. and An,G.
  TITLE     Isolation and characterization of an anther-specific gene, RA8,
            from rice (Oryza sativa L.)
  JOURNAL   Plant Mol. Biol. 39 (1), 35-44 (1999)
  MEDLINE   99178792
   PUBMED   10080707
REFERENCE   2  (bases 1 to 2757)
  AUTHORS   Jeon,J.-S., Chung,Y.-Y., Lee,S. and An,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JAN-1998) Life Science, Pohang University of Science
            and Technology, Pohang 790-784, Republic of Korea
FEATURES             Location/Qualifiers
     source          1..2757
                     /organism="Oryza sativa"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4530"
     TATA_signal     899..905
     mRNA            join(951..1042,1177..1309,1904..2666)
                     /clone="RA8"
     5'UTR           join(951..1042,1177..1309,1904..2523)
     CDS             join(1001..1042,1177..1309,1904..2523)
                     /note="RA8"
                     /codon_start=1
                     /product="anther-specific protein"
                     /product="anther-specific protein"
                     /protein_id="AAC99621.1"
                     /db_xref="GI:4091010"
                     /translation="MASLVAIAIAMALMVVQPGRQMTAFAARTSPAAAEAFWRAAMP
```

```
polyA_signal          2662..2666

ORIGIN

Query Match          75.2%;    Score 1079.8;  DB 8;  Length 2757;
Best Local Similarity 99.8%;   Pred. No. 8.6e-242;
Matches 1081; Conservative  0; Mismatches  2; Indels  0; Gaps  0;

GAPMPDAIVELLHHEHGVASAGKANGGDGPPPPMNFNYDDYRALPRSDAPSPDALN
RVAAVQNADENGVSSPPPPPTVFFPLEDAVRVGSSLPLPRPAADATAAGAAAAARPLPP
LRLTVRSVRAIESSFVVCRGETTAGVAGVYGCRHAATGSPARAVAVDAAGGGGGDAVI
AAVVCHADTSRWDPDHAAFRLLGVRPGGAAVCRAVADAHILPTNKD"

Qy      354 TCAGAATTCGAGGATCCGCCGTCGTTCGTTTTCCTGAAATCTGCAAGTCCCAGCAGCA 413
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      108 TCAGAATTCGAGGATCCGCCGTCGTTCGTTTTCCTGAAATCTGCAAGTCCCAGCAGCA 167

Qy      414 GCAGCAGCAGCAAGAGCAAATGCGGTGCGCCGGGCGCAGGGAGTTGATACTTTGANT 473
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      168 GCAGCAGCAGCAAGAGCAAATGCGGTGCGCCGGGCGCAGGGAGTTGATACTTTGANT 227

Qy      474 CTAGGGGTTCGTTCCATGTGGCTCTCACGCGCCGTGCGAATGTGCCATGACACTAGCTA 533
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      228 CTAGGCGTTCGTTCCATGTCGCTCTCACGCGCCGTGCGGAATGTGCCATGACACTAGCAT 287

Qy      534 CATCGGCCAAGATTATATTCCTCACATTTTTTCTTTCTTCCTATCGTCGTTCTGTTTGG 593
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      288 CATCGGCCAAGATTATAATTCCTCACATTTTTCTTTCTTCCTATCGTCGTCTGTTTGG 347

Qy      594 GAGCTTAAAATTATGAAAAGCAGCAGCTGCTGAGAAGCTAGCTGGTGGGAAATCTTGAAGAATTT 653
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      348 GAGCTTAAAATTATGAAAAGCAGCTGCTGAAAGCTAGCTGGTGGTGAGAATCTGAAGAATTT 407

Qy      654 GAGTTCTACGTTCAATTCTCCAGATTCTTACAATTACAGATTCTTATATAATTTAGGTAAAAAG 713
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      408 GAGTTCTACGTTCAATTCTCCAGATTCTTACAGATTCTTATAATTTAGGTAAAAAG 467

Qy      714 CTGGACTGTTTGGGAGCTTCTGTCGTCTCGGCCCGGAGATTCTGTGTGAGAAGCTGCAGCTGCTAGAA 773
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      468 CTGGACTGTTTGGGAGCTTCTGTCGTCACCGGAGATTCTGTGTGAGAAGCTGCAGCTGCTAGAA 527

Qy      774 GCTTCCCCAAACAGACCCCTAGTTGTACTCTAGCTGATCGATTCACTCTATTATTTATATAC 833
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      528 GCTTCCCCAAACAGACCCCTAGTTGTACTCTAGCTGATCGATTCACTCTATTTTATATAC 587

Qy      834 ACCTTGCTCTCTAGCTTATCAAAACGTAGCCAAGACTTGAATTTTAAAGCTTAAATTGATT 893
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      588 ACCTTGCTCTCTAGCTTATCAAAACGTAGCCAAGACTTGAATTTTAAAGCTTAAATTGATT 647

Qy      894 TTGATGTTCTTTCATCGTAATTCACTTACCGACCTTAGTTCGGCATTTGAATTTTTTAAAA 953
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      648 TTGATGTTCTTTCATCGTAATTCACTTACCGACCTTAGTTCGGCATTTGAATTTTTTAAA 707

Qy      954 ATAAATTTTTAGAGCTGATTTTGATTTTTTTTTCAGCGGAATTTATTTTCAGTATGTAAA 1013
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      708 ATAAATTTTTAGAGCTGATTTTTGATTTTTTTTCAGCGGAATTTATTTTCAGTATGTAAA 767

Qy     1014 AGTTTTACCTATAAATTATTATTTAATTTTCAGCCGGAGTAAGCAATTAGTGTTATGGGTTATAAT 1073
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      768 AGTTTTACCTATAAAATTATTTAATTTTCAGCCGGAGTAAGCAATTAGTGTTATGGGTTATAAT 827

Qy     1074 CATCTGGTATGCTTAAAATCTCTTTACTTGGACTTAGTTGGACAAATTCGTAAATGCATTCT 1133
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      828 CATCGGTATGCTTAAAATCTCTTTACTTGGACTTAGTTGGACAAATTCGTAAATGCATTCT 887

Qy     1134 CGTGGCCATCTCTTATAATACGGGCCGCTGCTTGCTCTTGTATCTGCACAAGAACTAG 1193
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      888 CGTGCCATCTCTTATAATACGGCCTGCTGCTTCGTCTTGTATCTGCACAAGAACTAG 947

Qy     1194 CTGGCAAAAGTCCTCAAGGCGGAACCCGCCTCCATCTTCTTCCTCCCCATGGCGGT 1253
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      948 CTGGCAAAAGTCCTCAAGGCGGAACCGCCTCCATCTTCTTCCTCCCCATGGCGGT 1007

Qy     1254 CCCTGCTTCGCCATGGCCATGGCTTCCATGTGCAGAGATATCCAGCATGCTGAA 1313
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1254 CCCTGCTTCGCCATGGCCATGGCTTCCATGTGCAGAGATATCCAGCATGCTGAA 1313
```

RESULT 7
BD242735

```
LOCUS       BD242735           240 bp    DNA    linear    PAT 17-JUL-2003
DEFINITION  DNA containing rice anther-specific gene and transgenic plants
            transformed thereby.
ACCESSION   BD242735
VERSION     BD242735.1  GI:33052505
KEYWORDS    JP 2002528125-A/9.
SOURCE      Oryza sativa
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 240)
  AUTHORS   An,G., Jeon,J.S., Chung,Y.Y. and Lee,S.
  TITLE     DNA containing rice anther-specific gene and transgenic plants
            transformed thereby
  JOURNAL   Patent: JP 2002528125-A 9 03-SEP-2002;
            SYNGENTA PARTICIPATIONS AG
COMMENT     OS    Oryza sativa (rice)
            PN    JP 2002528125-A/9
            PD    03-SEP-2002
            PF    02-NOV-1999 JP 200579761
            PR    03-NOV-1999  JP  1998/46973,19-NOV-1998 KR  1998/50126 PI
            GYNHEUNG AN,JONG SEONG JEON,YONG YOON CHUNG,SICHUL LEE PC
            C12N15/09,A01H5/00,C07K14/415,C12N15/00
            CC    DNA containing rice anther-specific gene
            CC    and transgenic plants
            CC    thereby      transformed
FEATURES             Location/Qualifiers
     source          1..240
                     /organism="Oryza sativa"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4530"
ORIGIN

Query Match          16.7%;    Score 240;  DB 6;  Length 240;
Best Local Similarity 100.0%;  Pred. No. 2.3e-45;
Matches 240; Conservative  0; Mismatches  0; Indels  0; Gaps  0;

Qy     1197 GCAAAGTCCTCAAGGCGGAACCGCCGCCTCCATCTTCTCCTTCCAGCTCCTCCCCATGGCGGTCTCC 1256
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 GCAAAGTCCTCAAGGCGGAACCGCCGCCTCCATCTTCTCCTTCCAGCTCCTCCCCATGGCGGTCTCC 60

Qy     1257 TCGTGCCATCGCCATGGCTCTCATGGTGCAGAGATATCCAGCATGCTGAAATT 1316
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 TCGTGCCCATCGCCATGGCTCTCATGGTGCAGAGATATCCAGCATGCTGAAATT 120

Qy     1317 AACTTGACGCATATCATCTCATCACTTGCATTCACTTCTGATTGTGCAGACATG 1376
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 AACTTGACGCATATCATCTCATCACTTGCATTCACTTCTGATTGTGCAGACATG 180

Qy     1377 TTTAGCTGATCAGTAAAACGTTGCCGTGTTGAATTGGTCCTTATCAGGTGGTTCAACCTGG 1436
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1377 TTTAGCTGATCAGTAAAACGTTGCCGTGTTGAATTGGTCCTTATCAGGTGGTTCAACCTGG 1436
```

```
Db     1008 CCCTGCTTCGCCATGGCCATGGCTCTCATGGTGGTCAGAGATATCCAGCATGCTGAA 1067
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     1314 ATTAACTTGACGCATATCATCTCATCATCACTTGCATTCAACTTCTGGATTGTGGATTGTGCAGAC 1373
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1068 ATTAACTTGACGCCATATCATCTCATCATCACTTGCATTCAACTTCTGGATTGTGGATTGTGCAGAC 1127

Qy     1374 ATGTTTAGCTGATCAGTAAAACGTTGCCGTGTTGAATTGGTCCTTATCAGGTTGGTTCAACC 1433
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1128 ATGTTTAGCTGATCAGTAAAACGTTGCCGTGTTGAATTGGTCCTTATCAGGTTGGTTCAACC 1187

Qy     1434 TCG 1436
            |||
Db     1188 TCG 1190
```

Db   181 TTTAGCTGATCAGTAAACGTTGCCGTGTGTTGAATTGGTCTTTATCAGGTGGTTCAACCTGG 240

RESULT 8
LOCUS   AK109240
DEFINITION   Oryza sativa (japonica cultivar-group) cDNA clone:002-156-F12, full
             insert sequence.
ACCESSION   AK109240
VERSION   AK109240.1 GI:32994449
KEYWORDS   FLI_CDNA; oligo capping.
SOURCE   Oryza sativa (japonica cultivar-group)
  ORGANISM   Oryza sativa (japonica cultivar-group)
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
             Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   The Rice Full-Length cDNA Consortium, National Institute of
             Agrobiological Sciences Rice Full-Length cDNA Project Team;
             Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
             Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
             Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
             Ohtsuki,K., Shishiki,T.,
             Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K.,
             and Genome Science Science Laboratory in Riken Genomic Sciences Center
             Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
             Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
             Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Nikura,J.,
             Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
             Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN;,
             Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
             Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
             Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
             Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
             Yoshino,M. and Hayashizaki,Y.
TITLE   Collection, mapping, and annotation of over 28,000 cDNA clones from
             japonica rice
JOURNAL   Science 301 (5631), 376-379 (2003)
MEDLINE   22752273
PUBMED   12869764
REFERENCE   2 (bases 1 to 598)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
             Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
             Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,R., Hiraoka,T.,
             Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,
             Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
             Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,
             Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
             Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
             Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,
             Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,
             Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,
             Namiki,T., Narikawa,R., Nishi,K., Nomura,K.,
             Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H.,
             Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saito,H., Sakai,C., Sakai,K.,
             Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K.,
             Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S.,
             Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,
             Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A.,
             Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W.,
             Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and
             Yoshimura,A.
TITLE   Direct Submission
JOURNAL   Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of
             Agrobiological Sciences, Department of Molecular Genetics, Head of
             Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
             305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
             Tel:81-29-838-7007, Fax:81-29-838-7007)
COMMENT   This clone is one of the 28K full-length cDNA clones from japonica
             rice.
             URL : http://cdna01.dna.affrc.go.jp/cDNA/
             NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K.,
             Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
             Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,
             Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and
             Yamamoto,M.
             FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y.,
             Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
             Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J.,
             Mizuno,K., Narikawa,R., Nikura,J., Oka,M., Ryu,R., Sugano,S.,
             Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,
             Yoshimura,A., Matsubara,K. and Murakami,K
             Genome Exploration Research Group in Riken Genomic Sciences Center
             and Genome Science Science Laboratory in Riken: Adachi,J., Aizawa,K.,
             Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T.,
             Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,R.,
             Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y.,
             Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,
             Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
             Koya,S., Kurihara,C., Matsuyama,T., Murata,M., Murata,M.,
             Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N.,
             Ota,Y., Saito,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H.,
             Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
             Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,
             Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K.,
             Yasunishi,A. and Hayashizaki,Y.
FEATURES   Location/Qualifiers
     source   1..598
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="002-156-F12"
ORIGIN

Query Match   12.9%;   Score 184.6; DB 8; Length 598;
Best Local Similarity   92.0%;   Pred. No. 2.2e-32;
Matches 240; Conservative   0; Mismatches   14; Indels   7; Gaps   4;

Qy   1182 CACAGAAACTAGCTGGCCAAAGTCCTCAAGGGGAACCGGCCTCCATCTTCTCCTTCAGCTC 1241
           ||||||||||| ||||| ||| | |||||||| |||| ||||| ||||||||||||||||
Db   1 CACAGAAACTAGCT-GCAAAGTCCTCAAGGGGAACGGCCTCCATCTTCTCCTTCAGCTC 59

Qy   1242 CTCCCATGGGCGGTTCCCTCGTCGGCCACGCCATCGCCATGGCTCTCATGGTGCAGAGATATC 1301
           |||||||| ||| |||||||| ||||||||||||||||||||||||||||||||||||| |
Db   60 CTCCCATGGGCGGTCCCCTCGTCGGCCATCGCCATCGCCATGGCTCTCATGGTGCAGAGGATAT 119

Qy   1302 C----AGCATGCTGAAATTAACTTGACGCATATCATCATCATCATCATCAATTCGAATTCAAACT 1357
           | |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   120 CCCAGCCATGGCTGAAATTAACCTGACGCATATCATCATCATCATCATCAACTTGCATTTCAAACT 179

Qy   1358 TCTGGATTGTGCAGACATGTTTAGCTGAT-CAGTAAACGTTGCCGTGTTGAATTGGTCTCCT 1416
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   180 TCTGGATTGTGCAGACATGTTTAGCTGATCGCAGTAAACGTTGCCGTGTTGAATTGGTCTCCT 239

Qy   1417 T-ATCAGGTGGTTCAACCTGG 1436
           | ||||||||||||||||||||
Db   240 TGATCAGGTGGTTCAACCTGG 260

RESULT 9
LOCUS   OSJN00063
DEFINITION   Oryza sativa genomic DNA, chromosome 4, BAC clone: OSJNBb0086G13,
             complete sequence.
ACCESSION   AL606706
VERSION   AL606706.3 GI:32480089
KEYWORDS   HTG.
SOURCE   Oryza sativa (japonica cultivar-group)
  ORGANISM   Oryza sativa (japonica cultivar-group)
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
             Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   Fu,G., Wang,S.Y., Ren,S.X., Lv,G., Lin,W., Gu,W.Q., Zhu,G.F.,
             Tu,Y.F., Jia,J., Yin,H.F., Zhang,Y., Cai,Z., Chen,J., Kang,H.,
             Chen,X.Y., Shao,Y., Sun,Y., Hu,Q.P., Zhang,X.L., Zhang,W.,

TITLE   Direct Submission
JOURNAL   Submitted (08-SEP-2001) Han Bin, National Center for Gene Research,
          Chinese Academy of Sciences, 500# Cao Bao Road, Shanghai 200233,
          CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:
          bhan@ncgr.ac.cn
REMARK   Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC
          clone: OSJNBb0086G13.
COMMENT   On Jul 8, 2003 this sequence version replaced gi:21741205.
          Web site: http://www.ncgr.ac.cn
          -------------------- Summary Statistics
          Assembly program: phrap

          This is a complete sequence.
          Genes were identified by a combination of several methods: Gene
          prediction programs including Fgenesh (http://www.softberry.com/),
          genscan (http://CCR-081.mit.edu/GENSCAN.html), GeneMarkHMM
          (http://genemark.biology.gatech.edu/GeneMark/), tRNAscan-SE (Sean
          Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/), searches of the
          complete sequence against NCBI none redundant protein database (nr)
          (ftp://ncbi.nlm.nih.gov/blast/db) and the EST database at NCGR.
FEATURES   Location/Qualifiers
     source   1..127952
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /sub_species="japonica"
                     /db_xref="taxon:39947"
                     /chromosome="4"
                     /clone="OSJNBb0086G13"
                     /clone_lib="CUGI-OSJNBb"
     gene   complement(2293..8874)
                     /gene="OSJNBb0086G13.1"
     CDS   complement(join(2293..2546,2589..2847,3547..3646,
                     3773..3859,3996..4234,4456..4541,4951..5074,5653..8874))
                     /gene="OSJNBb0086G13.1"
                     /codon_start=1
                     /protein_id="CAD41029.1"
                     /db_xref="GI:21741218"
                     /translation="MDAKDLVEIGMKEDITTMLFGKKVIELTEDAFDGSKERRKIFE
                     ELFCRTGTSGASTRHPRDGKSSSLRDASKELLSTNTPSSSASNHKSARCRVVESFTY
                     GNLSSYLVFCNNDKOKMOKOXTVGFPDMASASDLVIQWTPPSVERVTVRRAVARNRQRAK
                     LCSVLDLERVDITSVRQRDRDRGQHKYGHMASADLVLWDHLRLHANILMARLAKSGFX
                     DKSRVDHIYESPDKVMRLFSLPRAWKCFPCQWLLINSSGPGGNDYGRIWFNIHPFLSDL
                     KNTLCLEYEVRLPKQSLSFLNQWOLLDPFMAVVCIDKKVRALKKGVALKAVNSTVTP
                     VSCIESELSNPRNADNPLGLNHSNNFNSTHPRSRKNLLPLLQSDDLGNSSRSRKHPST
                     VGRKKSSQYKMDERLPIMAEIOEKSIRSTAHRIVMGLQDSTAFPGSRTSLNMKNKPSY
                     SKSNLTOSIQDKSDPLYFPPNYSPSDLVENAEINDPSFHAYDMIESAVDNSADSPDE
                     LILGANLLFSHEVDEMLLDDISNEHYCTAAVSEPQAKNIDLEDGPSGPSSVVSEKDTG
                     LKANKDVINNEOHDGVVAESOLANTDVWDRPAGALSFLPTMDTVLEGNEMSSBEMTK
                     GGQLSCEGTNNAMVISEPOLLFVSPHDGTLSFMNDSMCSQEMRSCLNASNGTMGTNVQ
                     LGIQSSVEASLIRGFLYLDNEGSPICWTVLNEPEPROLICSRPISKASELHGEMMME
                     CGMLTSGPKQTSKLDSSKNNRKRSKKAANIEDKGSQKOKVNDVPVSYCAIGQYMDHT
                     TENPAGCLISNDKEQIGSASTEQVSLNLLPGNSNDMNQQVPIKNSDNDQARVRIGSAE
                     KIVPEKTYKDVKROKASSRHKCKFDNDLLVTAVIHKLTARYRNRFSRRLAKKSGFX
                     RLPRCRWKSEERRDRKTFPKGARTVLGKLLEMGIVCKNILQYRPGSKNVLKDGNIT
                     KKGIRCRCCDMVFTMSMFXYHAGLROEIPSLNLFLGSGKSYTLCQLOAWSIEHKARKE
                     RAKCTMPLQADENDTCYGLCGDGGELICCDNCPASYHQDCTAGTDYCYVRCLC
                     DICEVINLKELRSSLPALECACQEROYHAKCIYGKLLCNEEGGPCAWFCGRRCOQIY
                     MNLRSRVGIFIHTDGFSCTVLRNNGDQRVSTAADIAILABCNMKLVIALSIMEECFL
                     PIIDARTGIDIIPPILVNWRSDFVHLDYGFYTVLNENDRILVSVASIRLHGFTVVAEM
                     PLTATCLENRQOGMCRRLMDYIEOMLKSLKVEMLLSAIPSLVDTWTMAFGFVPIDDL
                     DRKNLSRLRLVSVPGTVLLKRNLYECPGTNAGELRNPEPFKVISRSARKORDAGGGGG
                     GGGSKENLAFVDRSVAVTRADAGQORPFEDRSPVSCAVDSLADELRSLKINSSCENRR
                     RLGKTTVNSASTLVERPVGC"
     gene   13203..15107
                     /gene="OSJNBb0086G13.2"

                                                                    CDS   join(13203..13418,13855..14475,14535..14720,14793..15107)
                                                                                    /gene="OSJNBb0086G13.2"
                                                                                    /codon_start=1
                                                                                    /protein_id="CAD41028.1"
                                                                                    /db_xref="GI:21741217"
                                                                                    /translation="MANFPPMALKGQARGWLMNLPPVSVWSWEDLCCQOFTMNFQGTYP
                                                                                    RPDERADLHAVQRRDDESLRSYIQRFCQRTRKWEEERRQERREGKPAPASGNRSEA
                                                                                    KQKAPAEDVDDDDDLGFQEPGATIATVDGGACHASRSLKAMKRELLAAAPTHEAT
                                                                                    RRARWSEVALITDQTDDHPCVARGGQIAMVVSPTVCNMKLGRVLLIDGGAALNLSPAA
                                                                                    FDAIKVPGMVLRPSOPIIGVTPGHTWPLGHIDLPVTPGGSANFRTERVNFDVADLSLP
                                                                                    YNAMKMPGPGGPISVHGDLKVALACMEQRADHLAAASKPEGGNERLSTSVPAAPRQRM
                                                                                    ITCDEVEPDALVSFLRANADVPAWRPADMPGVPRKVIEHRLAVQTGARPVROKVRROA
                                                                                    PEROAPIRBEVARLLEAGFIREVIHPEWLANPIVVPKANGKLRMCIDYTDVTP"
                                                                    gene   18618..21317
                                                                                    /gene="OSJNBb0086G13.3"
                                                                    CDS   join(18618..19817,19926..21317)
                                                                                    /gene="OSJNBb0086G13.3"
                                                                                    /codon_start=1
                                                                                    /protein_id="CAD41027.1"
                                                                                    /db_xref="GI:21741216"
                                                                                    /db_xref="UniProt/TrEMBL:Q7XI15"
                                                                                    /translation="MAVLSRFISRLGEKTLPLFKLLKRSGPFIVTEEARRAFTOLKAY
                                                                                    LSSPPVLVAPDPDEPLLLYLABTPQVVSAALVVERDEDNPHPAHPHVSTRPGREQGG
                                                                                    EAPEPNGGRPPTAGAGPEHVCPTVPGAGPQODGPEAHLRAAAPRRLSPSDEVVUGEAEC
                                                                                    APRGLSDERRPGDAAPSEDRPERKKWQRPVYFVSEALRDAKTRVPQAOKMLYAILMAS
                                                                                    RKLRHYFQAHRVRTVTVTSYPLGQILHNREGTGRVVKVWAIELSEPDLHFPRHAIKSQAL
                                                                                    ADFVAEWTRAPEHVSAPEASLGPSQLPHFAYWMFDGSLSLLQCTGAGTLTSPSLDV
                                                                                    LRYMVRLDFRATNNMAEYEGLLAGLRVVAGLGIRRLLVLGDSQLVVNQVSKEYQCSDP
                                                                                    QMDAVRRQTPPGAFEERLAQPSVREDPLGETDAPERPPRPVEVQASGPEGSAPSSLRL
                                                                                    IAWITEIQAYLTDKTLPEDREGSRVORISKRYLVLVEETLYIRRAANGVLLKCIRREQG
                                                                                    VELLADIHHEGECRAHSASRTLVGKAFROGFYWPTALNDAVDLVRRCRACQFHAKQTHQ
                                                                                    PAQALQTIPLSWPFAVWGLDILGPFPRRAPGGFETLYVAVDKFTKWPBAYFVIKIDKHS
                                                                                    ALKFIRGITARFRVPNHIITDNGTQFTSELFGDYCKDMGIKLCFASPTPRSNGQVER
                                                                                    ANAEILKGLKTKTENILLKHGDSWIEELPAVLWANRTTPSRATGETPFILVYGAEAVL
                                                                                    PSELTLRSPRATMYCEADQDOLRRDLDYLEBRRRAALRAARYQQSLRCYHQRIHVRA
                                                                                    RSLCVDDLVLRRVQTRAGLSKLSPMWEGPYRVIGVPRPGCSVRLIATGDGTELPNPWNIE
                                                                                    HLRRPYP"
                                                                    gene   complement(22136..24278)
                                                                                    /gene="OSJNBb0086G13.4"
                                                                    CDS   complement(join(22136..22600,23718..24278))
                                                                                    /gene="OSJNBb0086G13.4"
                                                                                    /codon_start=1
                                                                                    /protein_id="CAD41026.1"
                                                                                    /db_xref="GI:21741215"
                                                                                    /db_xref="UniProt/TrEMBL:Q7KW5"
                                                                                    /translation="MMTGAGEEATRAAGGSCTERTACSTPLQCRRQRRRREMTSRRG
                                                                                    GTMWTTWTTNTGGVRSKGRRGRRGRPCCSDAGGDGDGDVGRRTGTAERAAEGDVDEEEGEAG
                                                                                    AGDGVPVKRSSMRRGQRCDVGBGVDEBPTEGEVDLADEEATPASPGGRRGGDGAGD
                                                                                    EGGVAEPRVVATSTGAQGAGRDRREVDTTLDQRAAEFEARAKELDARASRGGAAAGD
                                                                                    GDLAARLAAABHTIABLQGALDSSAGEVEALRLMGEVGPGMLRDAVSRLDRAGROAGL
                                                                                    WGGRTAKYAANOGGLAQRLSEMAGTIQRLPEELEETIKSSSRDLARGAVELVLASYQA
                                                                                    RDPDFSP"
                                                                    gene   complement(26701..28295)
                                                                                    /gene="OSJNBb0086G13.5"
                                                                    CDS   complement(join(26701..27871,28132..28295))
                                                                                    /gene="OSJNBb0086G13.5"
                                                                                    /codon_start=1
                                                                                    /protein_id="CAD41025.1"
                                                                                    /db_xref="GI:21741214"
                                                                                    /db_xref="UniProt/TrEMBL:Q7XYC3"
                                                                                    /translation="MARAGAAGGNNLPREGVVLLGPVTGARGAAALPTPCGRGNRRIP
                                                                                    PPHLIPEGAGREIPSRPSRHSKSGRSEAEETVTRRAARRRAADRRRAANCLRRAERAAQ
                                                                                    AQEAAEBARREARAEEAARERRAAARAAERGVRAEAAARLRRRAARRVRETTEAAHDE
                                                                                    AAGAPLGPDPDSGDVRDEPASGDARFBVGTSIGGPSRAASLPRRLSPTPSAAPLSAEPLL
                                                                                    QALAAANTTVLDGLSAQMVLQAERAELDAAWARVEEGRRSVEAMVEAGRKAHRRHAS
                                                                                    ELEARRDLAEIAREVEEERRETALIATAVNDTRADDLCLQYGRRAAAELEKKLDASRQR
                                                                                    ERALEARAATVLEELARTLDTKERRLVGRRAAVAIREATLAVHEAACABESSALRLRED
                                                                                    ELTRERRALKEAEVAAQLRREQARLNLEGARARARAALDVATGKTAIAFSRTLVY"
                                                                    gene   complement(34773..38123)
                                                                                    /gene="OSJNBb0086G13.6"
                                                                    CDS   complement(join(34773..35330,35438..35598,35691..35783,

gene   13203..15107

LOCUS   OSJN00050   156002 bp   DNA   linear   PLN 10-FEB-2004
DEFINITION Oryza sativa genomic DNA, chromosome 4, BAC clone: OSJNBa0088K19,
            complete sequence.
ACCESSION   AL606226
VERSION   AL606226.3 GI:32488551
KEYWORDS   HTG.
SOURCE   Oryza sativa (japonica cultivar-group)
ORGANISM   Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   Feng,Q., Zhang,Y., Hao,P., Wang,S., Fu,G., Huang,Y., Li,Y., Zhu,J.,
            Liu,Y., Hu,X., Jia,J., Yin,H.F., Zhang,Y., Zhao,Q., Ying,K., Yu,S.,
            Weng,Q., Zhang,L., Lu,Y., Mu,J., Lu,Y., Zhang,L.S., Yu,Z., Fan,D.,
            Liu,X., Lu,T., Li,C., Wu,Y., Sun,T., Lei,H., Li,T., Hu,H., Guan,J.,
            Wu,M., Zhang,R., Zhou,B., Chen,Z., Chen,L., Jin,Z., Wang,R.,
            Yin,H., Cai,Z., Ren,S., Lv,G., Gu,W., Zhu,G., Tu,Y., Jia,J.,
            Zhang,Y., Chen,J., Kang,H., Chen,X., Shao,C., Sun,Y., Hu,Q.,
            Zhang,X., Zhang,W., Wang,L., Ding,C., Sheng,H., Gu,J., Chen,S.,
            Ni,L., Zhu,F., Chen,W., Lan,L., Lai,Y., Cheng,Z., Gu,M., Jiang,J.,
            Li,J., Hong,G., Xue,Y., and Han,B.
  TITLE   Sequence and analysis of rice chromosome 4
  JOURNAL   Nature 420 (6913), 316-320 (2002)
   PUBMED   12447439
  MEDLINE   22337377
REFERENCE   2
  AUTHORS   Fu,G., Wang,S.Y., Ren,S.X., Lv,G., Lin,W., Gu,W.Q., Zhu,G.F.,
            Tu,Y.F., Jia,J., Yin,H.F., Zhang,Y., Cai,Z., Chen,J., Kang,H.,
            Chen,X.Y., Shao,Y., Sun,Y., Hu,Q.P., Zhang,X.L., Zhang,W.,
            Wang,L.J., Ding,C.W., Sheng,H.H., Gu,J.L., Chen,S.T., Ni,L.,
            Zhu,F.H., Han,B., Feng,Q., Huang,Y.C., Li,Y., Chen,L., Fan,D.L.,
            Hu,X., Liu,Y.L., Mu,J., Yu,S.L., Chen,Y.H., Lu,T.T., Zhang,Y.,
            Zhang,L., Lu,Y.Q., Yu,Z., Li,C., Liu,X.H., Lu,T.T., Zhang,Y.J.,
            Li,C., Li,T., Zhang,Y., Hu,H., Jia,P.X., Qian,Y.M., Ying,K.,
            Zhou,B., Chen,Z.H., Hao,P., Zhang,L., Wu,M., Zhang,R.Q., Guan,J.P.
            and Hong,G.F.
  TITLE   Direct Submission
  JOURNAL   Submitted (08-SEP-2001) Han Bin, National Center for Gene Research,
            Chinese Academy of Sciences, 500# Cao Bao Road, Shanghai 200233,
            CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:
            bhan@ncgr.ac.cn
REMARK   Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC
            clone: OSJNBa0088K19.
COMMENT   On Jul 9, 2003 this sequence version replaced gi:21740748.
            Web site: http://www.ncgr.ac.cn
            --------------- Summary Statistics ----------------
            Assembly program: phrap

            This is a complete sequence.
            Genes were identified by a combination of several methods: Gene
            prediction programs including Fgenesh (http://www.softberry.com/),
            genscan (http://CCR-081.mit.edu/GENSCAN.html), GeneMarkHMM
            (http://genemark.biology.gatech.edu/GeneMark/), tRNAscan-SE (Seen
            Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/), searches of the
            complete sequence against NCBI none redundant protein database (nr)
            (ftp://ncbi.nlm.nih.gov/blast/db) and the EST database at NCGR.
FEATURES   Location/Qualifiers
     source   1..156002
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /sub_species="japonica"
                     /db_xref="taxon:39947"
                     /chromosome="4"
                     /clone_lib="CUGI-OSJNBa"
                     /clone="OSJNBa0088K19"
     gene   complement(3548..4860)
                     /gene="OSJNBa0088K19.1"
     CDS   complement(join(3548..4022,4229..4860))
                     /gene="OSJNBa0088K19.1"
                     /codon_start=1
                     /protein_id="CAD40919.1"
                     /db_xref="GI:21740758"
                     /db_xref="UniProt/TrEMBL:Q7XBU3"
                     /translation="MPTATGRSRTPVRSASAVIAGTESGQGHHLKIDGYSRIKDELPTGS
                     DIKSRSFRAGGHSWHLRYYPNGRNSDCABCISIFLQLDYNVMKQVKAQYKFSLLDRAR
                     KPSYSRSSGKADVFLNTGWGYRTYIERGLLESSBEYLRDDCLTIVCDFTVFKDLRTEDI
                     DVDDAMPPQSPPTVVPSDLRHHLGGLLATGGEGADVTFEVEGKTFAAHRWVLAARS
                     PVRRVALFGATTGQADDVVRVNIDAMKVQDFRALLHYMYTDSLPEMKGGEAAAMLPDL
                     VAAANRYKMEIRLVCEHKLCEYVNGRTVVAMLAFAGEHQCNGLKEKCLRFLDDPVKL
                     RLIVQABGVENLSKSYPSILKDVIAKFVATPVDS"
     gene   10596..13605

CDS
    /gene="OSJNBa0088K19.2"
    join(10596..11575,11670..13605)
    /gene="OSJNBa0088K19.2"
    /codon_start=1
    /protein_id="CAB03206.1"
    /db_xref="GI:3248852"
    /db_xref="UniProt/TrEMBL:Q7XTK2"
    /translation="MTIAIFAHLCEMFIGVRPSLRLFRWFFTVQPVSPSVVGGCYFQ
    PRGPVLNRYIPCVLRKKWDDWKSDWFYTPLADEARLRPLSQPPAQASWRAPVDLGDG
    YDAVLNRLAGLRSQGLTGAMVYGDYLVRRIAPLQRRARGAWEYTGSEDCMGTHRGVKW
    DWVPEFKWMIQRVLNLSSVEASLIPQGVLPLCSDPDRANILITNMWAVGASEERAPKG
    LDGASGSRKDQSTPGGGRASGPRDGGSGGCRPADARGRKRKQEGTSSPSSPPRGGGVAR
    VSSRRPBGAAPTPQPGERKKKRLRKMGGTHBPSRGNLISPPRWRSFNRPPRSKIPSRPS
    CQSKSGQSRAEAENAATTEARRESDREAAADRLREAEEAAARARARAQAEEAAAREBARI
    RQAERATTSSEATRETDRQEAADRLREAEEAAARARARAQAEEAAAREARL
    SGAARDEAAGASLGPTPSGDTQDKRGPGDVPEPCTSIGGPSRAASSLRRLSPTTSAAP
    LSAEPLLQALAVANTVLGIARVEBERGALIATTVMNRAQDDLRLQYGSWEAELGKK
    AHRHASELEARRKVLGEIAREVEBERGAALIATTVMNRAQDDLRLQYGSWEAELGKK
    LDAARGVLDAAAAREBEVASRRREEALERARAMALEBRACVVERDLADREAAVA
    IREVTLAHEAEACAEEESALRLEEDTLTEREEALERARAMALEBRACVVERDLADREAAVA
    ARNLBGARAERAALNQRAAELEAQARELDARASGGAATGDSDLAARLAAAEHTIAD
    LQGVLNSSTGEBVEALRLAGBIGPSMLRDAVSRLDRAGRQGAGLWGGRRTVKYAANQGLA
    QRLSEMAGTLQRLPEELEGTIKSSEAGSAPRLAFALNSDEBDDDGGVGDSGDEAGDPGASE"

gene
    complement(14232..20388)
    /gene="OSJNBa0088K19.3"

CDS
    complement(join(14232..15623,15786..17157,17234..17934,
    18034..19130,19302..20388))
    /gene="OSJNBa0088K19.3"
    /codon_start=1
    /protein_id="CAD40917.1"
    /db_xref="GI:21740756"
    /db_xref="UniProt/TrEMBL:Q7XU22"
    /translation="MAEQKAHNLSPSVSGDDGENPRRRARTPPPPRQSPRREEAL
    ERVEGSATSTSTGDGEGRRDGERHLVYGDGSTPGGALQARGALLRHPVVDPESPA
    QRWLDDVANLVMTARQRLGAGGRSSATKTSGAATTGSVSSRRARRAAAVARHSAATP
    SSAPPTREDLRGGPDARISIERRRNGRRAAHATEGASSSGVSPQHGRGDQPSVPPVGG
    VGCRAFVASLRRVDRWFRRPTIAEKYDGSVNPABFLQVTTGIBAAGGDRVNANFF
    PMALKQQARGWLMNLPPASVHSWEDLCCQFTMNPQGTYPRPGEEADLHAVQRRDDESL
    RSYIQRFCQVRNTIPCIPAHAVIYAFRGGRRDRRRKKRSTHSDDEGHVLAVEGASRAI
    RKGRPASDKKKEASASRERPTGKWCTVKNTSLHDLADCRAVKSLAERTRKWEEERRQ
    ERREGKSPAVPSGNRRSEAKQKAPAEDIDDGDDLGFQRGPATIATVDGCAHVSRR
    SLRAMKRELLAAAPAHEATRRARWSEFALTPDCTDHPPCVARGGIAMVVSPTVCNVK
    LGRVLIDGGAALNISPAAFDAIRAPGWHIRSQPIIGVTGTWHFPLGHIDLPVTFGG
    SANRTERVRNPDVADLSLPYNAVLGRPALVKFMPAVHVAYLQMKMPGPGGPISVHGDL
    KVALACMEQRAGHLAAASKPEGGDERLGTSVPAAPRORMITCDEVPANADVFAWRPAD
    MPGVPREVIEHRLAVRPGAKPVRQKVRQKGFIREVARLLBAGFIREVHPEW
    LANPVVVPKANGKLQMCIDYTDLNKACPKDYPLPRIDQIVDSTAGCDLLCFLDAYSG
    YHQIRMAREBEEKTAFIPIATVCYTMWPGLKNAGPFPQRTTRISLGSQIGRNVEAY
    VDDLVVKFDRETLLSDLAEFTFESLRARIKLNPDKCVFGVPAGKLLGRLDVQCVTG
    CMALNRFISRLGEKALPLFKLLRRSGPFTWTEEARALTOLKAYLSSPPVLIAPEPD
    EPLLLYLAATPQVVSAALVVERDEDNPHSAHPHPVLAWPGREQGEBAPEPNGGPRPT
    TGAGPLPACQTVPGAPDQPGDGPGATAGRPHLSPSDPEVNPVFPRPGREQGEBAPENG
    GLRLPTTCVGPLPACPTTPGAPDPODGPEATVGRPHLSSSDPEVVGTEDKCAPRGRLD
    KERPRDAAPGEBDPHRKVQRPVYFVVSRALRDAKTRYPQAQKMLYAILMASRKLRHYF
    QAHRVTVVTSYPLAQILNRREGTGRVKVNAIELESFDLHPEPHHAIKSQALADPAEW
    TPAEPVSVPEASSGPSQLPHTARWVMQPDGSLSLQGAGAGVTLTSPSGDILRYLVRL
    DFRATNNMAREYEGLLAGLRVAAGLGIRRLLVLGDSQLVVNQTPGAFEBRLAQPSARP
    DFLGETDAPDRPRPVRVQASGPEGSAPSSLRIIAWIAEIQAAVITDKTLPEDREGSER
    VQRISKRVLIEGTLYRAANGILLKCIPREGIELLADIHEGECGAHSASRTLVGKA
    FRQGPYWPTALANDAVDLVRRCRACQFHAKQIHQPAQALQIIPLSWPAVWGDILGPF
    RRAPGGFRYLIVAIDKFTKWPEAYPVVKIKHSALKFIKGITARFGVPNRIITDNGTQ
    FISELFGDYCEDMGIKLCPASPAHPRSNGQVERANAEILRGLKTKTFNILKHGDSWI
    EELPAVWANRTTPSRATGETPPFLVGAEAVLPSELTLRSPRATMYCEADDQOLRRD
    DLDYLEERRRAALRAARYQOSLRRYHQRHVRARSLCVDDLVRCVQTCAGLSKLSPM
    WEGPYRVIGVPRPGSVRLATGDGIELPNPWNIEHLRRYP"

gene
    complement(22227..23261)
    /gene="OSJNBa0088K19.4"

CDS
    complement(22227..23261)
    /gene="OSJNBa0088K19.4"
    /codon_start=1
    /protein_id="CAD40916.2"
    /db_xref="GI:38345365"

gene
    complement(36584..37699)
    /gene="OSJNBa0088K19.5"

CDS
    complement(36584..37699)
    /gene="OSJNBa0088K19.5"
    /codon_start=1
    /protein_id="CAB03207.2"
    /db_xref="GI:38345366"
    /db_xref="UniProt/TrEMBL:Q7XQB8"
    /translation="MPTDAAAAAAASSPLLPAPPSAARRRRRRLLTSPNPSVSST
    GNCVQSRHFRVGGHDWIYRYYPNGFNSNVSDCISIYLVLDGHEAHDYYGRSIVRABL
    TLSLLDQEREPVTSIYIYSHGLQIFPDGYGRYRGSLRPIQKAVLERSEYLRDNRFTIRCD
    RSPVFHQELFSLTEKGNATGGAGVIIRVDDMEAQDFEALLHPIYTDSLPEMKGGDAV
    AMLPDLVAAANRYKMERLRLVCEBDLCEYTVRTVAAMLAPAGEHCPLEBKKCLQLL
    EDPANLRNIVETBGLEHLTKSYPFVLKDILIAMFATKP"

gene
    complement(55963..56634)
    /gene="OSJNBa0088K19.6"

CDS
    complement(55963..56634)
    /gene="OSJNBa0088K19.6"
    /codon_start=1
    /protein_id="CAB03208.2"
    /db_xref="GI:38345367"
    /db_xref="UniProt/TrEMBL:Q7XQB7"
    /translation="MPTDAAAAAAASSPLLPAPPSAARRRRRRLLTSPNPSVSST
    STSSSSSSSSSSSSSLSPPAPPSEPAPSSPFHHRFLSPLRASAVPPSWEHRPGIPKTPA
    RQHQHHRGGCGCGKVYSTPLPLPPESLLSSKVVVADEDRGGADRLAVSDDDDAKAARR
    RSRRRHRRLLQRPRRPAALAAALTDWLAVLSLYRSCTRSVDCLADAGQPPTTTTTP
    AAKAG"

gene
    complement(70247..73438)
    /gene="OSJNBa0088K19.7"

CDS
    complement(join(70247..70466,70571..70612,70735..70811,
    70887..71003,71238..71341,71419..71488,71578..71687,
    71934..72021,72095..72185,72265..72357,72451..72776,
    73011..73321,73426..73438))
    /gene="OSJNBa0088K19.7"
    /codon_start=1

/db_xref="UniProt/TrEMBL:Q7XU23"
    /translation="MPTSTSSRKPVRSASTIIAGTESGGQHLLKIDGYSHTKDKLPTPG
    SNVKSRSFRVGGHSWHISYYPSGNDSDKANCISIFLNLDDDVDVRKAQFKFSLLDRAGR
    QPARLQKQRVGVCEFHPEGRAGEIGVSPGRLPHDRVRPHRLHGAADGHRRRHRNAAA
    AAAYGGGPTVGPAPAPRRPPRHRGGRRRDVRGQREDVRDGTVOAGRGPGVGSVPCRAFW
    AKQGAGCHDDRRGGSHRHTHRRHGGAGLRGSAPLHVHRLAAGARDDERRRGRGGDAPG
    PGRGEBQVQDGEABAGVRAQTVRVRERENGCFHACIRPGAPLRWPQGGVLAFPRRSSE
    SERDREGSII"

gene
    complement(149728..149789)
    /gene="OSJNBa0088K19.7"

RESULT 11
AP005106/c
LOCUS   AP005106   126544 bp   DNA   linear   PLN 22-JUL-2004
DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA,
            BAC clone:OSJNBb0091119.
ACCESSION   AP005106
VERSION   AP005106.6 GI:50510301

Query Match 9.6%; Score 137.8; DB 8; Length 156002;
Best Local Similarity 82.1%; Pred. No. 2.9e-21;
Matches 183; Conservative 0; Mismatches 37; Indels 3; Gaps 2;

Qy   585 TCTGTTTGGGAGCTTAAAATTATGAAAAAGCAGCTGCTGAGAAAGCTAGCTGTGAGAATCT 644
             ||||||||||| ||||||| |||||||||||||||||||||||||||||||||||||||
Db   149907 TGTTTGGGAGAGCTTAAGATTCTGAGAAGCAGCTGCTGAGAAAGCTAGCTGTGAGAATCT 149848

Qy   645 GAAGAATTTGAGTTCTCACGTTCATTCTCCAGATTCTACAATTACAGATTCTTATAATTTA 704
             ||||||| ||||| |||||||||| || ||||| |||||| |||| ||||||||||||
Db   149847 GGAGAAGCTGCTTCTA--GTTCATTTTCCAGATTTTACAACTACACAGTTCTTATAATTTA 149789

Qy   705 GGTAAAAAGCTGGACTGTTT---GGGAGCTTCGTGCTGAGAGATTCTGTGAGAAGCTGC 762
             |||||||||||||| |||| |||||||| |||||||| |||||||||||||||||||
Db   149788 GGTAAAAAGCTGGACTGTTTGGGGGGAGCTTCTTGGCTGAGGGATTCTAGAGGAAGCTGC 149729

Qy   763 AGCTGCTAGAAGCTTCCCCAAACAGACCCCTTAGTTGTACTCTA 805
             ||||||||||||||| ||||||||||||| |||||||| |||||
Db   149728 AGCTGCTAGAAGCTCCCCCAAACAGCCCAGACTCTGTTCCCAA 149686

```
KEYWORDS     .
SOURCE       Oryza sativa (japonica cultivar-group)
  ORGANISM   Oryza sativa (japonica cultivar-group)
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
             Ehrhartoideae; Oryzeae; Oryza.
REFERENCE    1  (bases 1 to 126544)
  AUTHORS    Sasaki,T., Matsumoto,T. and Katayose,Y.
  TITLE      Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, BAC
             clone:OSJNBb0091119
  JOURNAL    Published Only in Database
REFERENCE    2  (bases 1 to 126544)
  AUTHORS    Sasaki,T., Matsumoto,T. and Katayose,Y.
  TITLE      Direct Submission
  JOURNAL    Submitted (18-APR-2002) Takuji Sasaki, National Institute of
             Agrobiological Sciences, Rice Genome Research Program; Kannondai
             2-1-2, Tsukuba, Ibaraki 305-8602, Japan
             (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
             Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT      On Jul 22, 2004 this sequence version replaced gi:34495386.
             Genes were predicted from the integrated results of the following:
             GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), PGENESH
             (http://www.softberry.com/), GeneMark.hmm
             (http://opal.biology.gatech.edu/GeneMark/), GlimmerM
             (http://www.tigr.org/tdb/glimmerm/glmr_form.html), RiceHMM
             (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor
             (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4
             (http://globin.cse.psu.edu/html/docs/sim4.html), Gap2
             (http://www.tigr.org/software/glimmerm/), BLASTN and BLASTX. The
             genomic sequence was searched against NCBI NonRedundant Protein
             database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
             sequence database at RGP or DDBJ. Protein homologies of the coding
             regions were searched against NCBI NonRedundant Protein database
             with BLASTP. ESTs represent the identified cDNA sequences using
             BLASTN with the corresponding DDBJ accession no. and RGP clone ID.
             Full-length cDNAs represent the identified cDNA sequences using
             BLASTN with the corresponding DDBJ accession no.
             A gene with identity or significant homology to a protein is
             classified based on the protein name to indicate the homology level
             such as same name, 'putative' and '-like protein'. A gene without
             significant homology to any protein but with full-length cDNA or
             EST homology (covering almost the entire length of partial
             sequence) is classified as an 'unknown' protein. A gene predicted
             by two or more gene prediction programs is classified as a
             'hypothetical' protein according to IRGSP standard. A gene
             predicted by a single gene prediction program is also classified as
             a probable 'hypothetical' protein and is included as a
             miscellaneous feature of the sequence.
             The orientation of the sequence is from M13rev to -21M13 of the BAC
             clone. This sequence of OSJNBb0091119 clone has an overlap with
             OSJNBa0028021 (DDBJ: AP005766) clone at 5' end and with P0045F02
             (DDBJ: AP004268) clone at 3' end. Detailed information on overlap
             and assembly quality together with annotation of this entry is
             available at http://rgp.dna.affrc.go.jp/GenomeSeq.html.
FEATURES             Location/Qualifiers
     source          1..126544
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="7"
     gene            493..1428
                     /clone="OSJNBb0091119"
     misc_feature    493..1428
                     /gene="OSJNBb0091119.101"
     gene            1669..8396
                     /gene="OSJNBb0091119.101"
                     /note="retrotransposon protein-like"
     gene            1669..8396
                     /gene="OSJNBb0091119.102"
     misc_feature    1669..8396
                     /gene="OSJNBb0091119.102"
                     /note="probably inactive due to 3' exon missing in CDS
                     probably inactive due to including stop codon(s) in CDS
```

```
     gene            complement(join(8745..8943,9026..9099,9882..9936,
                     10764..10899,12192..12432))
                     /gene="OSJNBb0091119.103"
                     /note="hypothetical ORF
                     predicted by GENSCAN
                     this category is not included in IRGSP standard"
     misc_feature    complement(join(8745..8943,9026..9099,9882..9936,
                     10764..10899,12192..12432))
                     /gene="OSJNBb0091119.103"
                     /note="hypothetical ORF"
     gene            complement(13427..13621)
                     /gene="OSJNBb0091119.104"
     misc_feature    complement(13427..13621)
                     /gene="OSJNBb0091119.104"
                     /note="hypothetical ORF
                     predicted by GlimmerM
                     this category is not included in IRGSP standard"
     gene            complement(join(13790..14494,
                     14056,14141..14239,14405..>14494))
                     /gene="OSJNBb0091119.105"
     mRNA            complement(join(<13790..14056,14141..14239,14405..>14494))
                     /gene="OSJNBb0091119.105"
                     /product="hypothetical protein"
     CDS             complement(join(13790..14056,14141..14239,14405..14494))
                     /gene="OSJNBb0091119.105"
                     /note="start and end point are not identified"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAC22419.1"
                     /db_xref="GI:24414176"
                     /translation="MCSWFAACVRDKNILAHLREIAIALLSFYLKALNDTVGVGNITIF
                     EPRNEDMERDLDLEFRRHHPAPPLLPSTRGDTLVGIHOPIENDTSGGGGEAPTAKVPH
                     SDLGYEVAISLDTPLITFLPDESSSEEETVATWRLLITVVIDQFQPSILFK"
     gene            18215..19057
                     /gene="OSJNBb0091119.106"
     mRNA            <18215..>19057
                     /gene="OSJNBb0091119.106"
                     /note="start and end point are not identified"
     CDS             18215..19057
                     /gene="OSJNBb0091119.106"
                     /note="predicted by GeneMark.hmm etc."
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAC22418.1"
                     /db_xref="GI:24414175"
                     /translation="MYISPYSPQPPQNNLERKIQIRRMLQVQGSLRVHTFSSAAIVHA
                     VETSDEDSDPSHFCAVPQNGKA3RKEIRRIKRLLSSLGQKHHISKYFFR3RSBAAN
                     SNAVIDNRGGCGSDMETFVSAKSSELCSFRTDDDDSESRSFRLSPLPIPPTGIEFQP
                     PASPVKIIKKLPGYIIGRQLDGAPAAAAPSTKLSLSFKKLMRLVDKSKSKOMIKKKV
                     LRALKGRFGGGEBRGRDGHVREGKESSDYGDGDGDGDBDVFWRKDVRGLRCRRVEBDN
                     DLPY"
     gene            20513..20770
                     /gene="OSJNBb0091119.107"
     mRNA            join(<20513..20591,20694..>20770)
                     /gene="OSJNBb0091119.107"
     CDS             join(20513..20591,20694..20770)
                     /gene="OSJNBb0091119.107"
                     /note="start and end point are not identified"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAC22417.1"
                     /db_xref="GI:24414174"
                     /translation="MPSAATAVVLVHLGYGPHPRLPLPPSSFSSSTLAAASAAIVQRI
                     RVQSSSF"
     gene            complement(23334..23489)
                     /gene="OSJNBb0091119.108"
     misc_feature    complement(23334..23489)
                     /gene="OSJNBb0091119.108"
                     /note="hypothetical ORF
                     predicted by GlimmerM
                     this category is not included in IRGSP standard"
     gene            complement(25089..25464)
```

pseudogene, polyprotein"
/pseudo

RESULT 12
AC144738/c

```
LOCUS       AC144738              156649 bp    DNA     linear   PLN 22-JUL-2004
DEFINITION  Oryza sativa (japonica cultivar-group) chromosome 5 clone
            OSJNBa0029B02, complete sequence.
ACCESSION   AC144738
VERSION     AC144738.2  GI:40385890
KEYWORDS    HTG.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 156649)
  AUTHORS   Chow,T.-Y., Hsing,Y.-I.C., Chen,C.-S., Chen,H.-H., Liu,S.-M.,
            Chao,Y.-T., Chang,S.-J., Chen,H.-C., Chen,S.-K., Chen,T.-R.,
            Chen,Y.-L., Cheng,C.-H., Chung,C.-I., Han,S.-Y., Hsiao,S.-H.,
            Hsiung,J.-N., Hsu,C.-H., Huang,J.-J., Kau,P.-I., Lee,M.-C.,
            Leu,H.-L., Li,Y.-F., Lin,S.-J., Lin,Y.-C., Wu,S.-W., Yu,C.-Y.,
            Yu,S.-W., Wu,H.-P. and Shaw,J.-F.
  TITLE     Oryza sativa BAC OSJNBa0029B02 genomic sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 156649)
  AUTHORS   Chow,T.-Y. and Hsing,Y.-I.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-DEC-2003) Institute of Botany, Academia Sinica, 128,
            Section 2, Academia Road, Nankang, Taipei 11529, Taiwan
REFERENCE   3  (bases 1 to 156649)
  AUTHORS   Chow,T.-Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAY-2003) Institute of Botany, Academia Sinica, 128,
            Section 2, Academia Road, Nankang, Taipei 11529, Taiwan
REFERENCE   4  (bases 1 to 156649)
  AUTHORS   Chow,T.-Y. and Hsing,Y.-I.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-JUL-2004) Institute of Botany, Academia Sinica, 128,
            Section 2, Academia Road, Nankang, Taipei 11529, Taiwan
COMMENT     On Dec 30, 2003 this sequence version replaced gi:30725893.
            Genes were predicated from the integrated results of the following:
            BLASTN2.0, BLASTX2.0, GENSCAN (Chris Burge,
            http://genes.mit.edu/GENSCAN.html), Pgenesh
            (http://www.softberry.com/), GlimmerR
            (http://www.tigr.org/softlab/glimmer/glimmer.html), TWINSCAN
            (http://genes.cs.wustl.edu/) and GeneSplicer
            (http://www.tigr.org/tdb/GeneSplicer/index.shtml). The sequence was
            searched against the Swiss-Prot+TrEMBL protein database, the NCBI
            plant EST database, the TIGR Rice Gene Index and the rice
            full-length cDNA database (KOME,
            http://cdna01.dna.affrc.go.jp/cDNA/). Annotated genes are named to
            indicate the level of evidence for their annotation. Genes with
            similarity to other proteins are named after the database hits.
            Genes without significant peptide similarity but with EST
            similarity are named as unknown proteins. Genes without protein or
            EST similarity, that are predicted by more than two gene prediction
            programs over most of their length are annotated as hypothetical
            proteins. This clone overlaps with P0617A08(accession # AC135426)
            and P0478P09(accession # AC130610).
FEATURES             Location/Qualifiers
     source          1..156649
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="5"
                     /clone="OSJNBa0029B02"
     gene            complement(2649..8418)
                     /gene="OSJNBa0029B02.1"
                     /note="hypothetical protein"
     CDS             complement(join(2649..3247,7353..8418))
                     /gene="OSJNBa0029B02.1"
                     /note="hypothetical protein"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAT77340.1"
                     /db_xref="GI:50511417"
                     /translation="MAEEDSRVWIPPAABKLTDDLLVEVLSRVPYKSLCRSKRVCRR
```

```
     mRNA            complement(join(<25089..25196,25336..>25464))
                     /gene="OSJNBb0091119.109"
                     /note="start and end point are not identified"
     CDS             complement(join(25089..25196,25336..25464))
                     /gene="OSJNBb0091119.109"
                     /note="predicted by GeneMark.hmm etc."
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAC22416.1"
                     /db_xref="GI:24414173"
                     /translation="MKATSPLLTSEDELPAVSRNGNGGEAGBEEVAAKRWAMLGSEEA
                     GVBEADRGGEBENGRRVRRAAINGGSSLETKFDD"
     gene            28518..28863
                     /gene="OSJNBb0091119.110"
     mRNA            join(<28518..28609,28746..>28863)
                     /gene="OSJNBb0091119.110"
                     /note="start and end point are not identified"
     CDS             join(28518..28609,28746..28863)
                     /gene="OSJNBb0091119.110"
                     /note="predicted by GeneMark.hmm etc."
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAC22414.1"
                     /db_xref="GI:24414171"
                     /translation="MEEKLFELADTVHKYDDARSNGNGHLQQRGRRLAVVIIDGNFDAD
                     SSIEPLIGNEEVDPNKVVTVSARPK"
     gene            29573..35485
                     /gene="OSJNBb0091119.111"
     mRNA            join(<29573..29674,35129..>35485)
                     /gene="OSJNBb0091119.111"
                     /note="start and end point are not identified"
     CDS             join(29573..29674,35129..35485)
                     /gene="OSJNBb0091119.111"
                     /note="predicted by GeneMark.hmm etc."
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAC22413.1"
                     /db_xref="GI:24414170"
                     /translation="MPRGREROQHDASPILCRQAGKAIPKPSTPRWLDEYQOAAEGAQV
                     LRAGRHNRRNWGGGGSDDDGCTPAATVAITGIMAARREVIGEAKDDAVGDGBADAAAVA
                     VAGQGTAYRGCWGGGGRSGGGOGGODGATCVRDRALDHRCRVHPVGGSDDDG"
     gene            complement(36502..36711)
                     /gene="OSJNBb0091119.112"
     misc_feature    complement(36502..36711)
                     /gene="OSJNBb0091119.112"
                     /note="hypothetical ORF
                     predicted by GlimmerM
                     this category is not included in IRGSP standard"
```

```
Query Match       79.6%;  Score 127.4;  DB 8;  Length 126544;
Best Local Similarity   8.9%;
Matches 179;  Conservative   0;  Mismatches 36;  Indels 10;  Gaps   2;

Qy    576 TCCTAGTCGTCGTTCGTTTGGGAGCTTAAAAATTATGAAAAGCAGCTGCTGCTGAGGAGCTAGCTGG 635
          ||||| ||||||||||||||||||||||||||||||||        |||||||||||||||||||||
Db  72464 TCCTAGGGCCAGTTTGGGGAGCTTAAAAATTCTGAGAAGCAGCTGCTGAGGAGCTAGCTGG 72405

Qy    636 TGAGAATCTGAAGAATTGAGTTCTACGTTCATTCTCCAGATTCTACAATTACAGATTCT 695
          ||||||||       |||||||||     |||||||||||||||||||    |||||||
Db  72404 TGAGAATCTTCGTCGTTCTA------GTTCATTTTTCAGATTCTACAACTACAGATTCT 72352

Qy    696 TATATTTTAGGTAAAAAAGCTCGGACTCGTTT---GGGAGCTTCTGTCGTCCGCGAGATTCTGT 752
          |||  || ||||||||||||||||       |||||||||||||            ||||||
Db  72351 CCGAATTTGGGTAAAAAAGCTGGATTGTTTGGGGGAGCTTCTACGAACTACAGATTCTAG 72292

Qy    753 GAGAAGCTGCAGCTGCTAGAAGCTCCCCAAACACAGACCCCTAGTT 797
          |||||||||||||||||||||||||||||||||||||     |||||
Db  72291 GAGAAGCTGCAGCTGCTAGAAGCTCCCCAAACACAGCTCTTAGTT 72247
```

WRRVISDPDHRRLLPRYHLHDAVVGFDHRSFSFSFSGRGRFVGPSLPLPLPDCGDLRLL
DICNGLLLCRRRRLSDPRRFDYLVANPATEQWTLVLPESGWTHKBQIARLGFDPAVSS
HFHVFPFRLVEHSAMGMSGDHDGNVAVBIYSSETGVWIHRNNGWGCIIRTLDIWRSV
FPKGMLHLITMDDVAVVDVSSBTGVWIHPEITFQPDYIQVDDGFIDVSQDCLCFVNT
DRDDLYKLSVVWLEDYSSDQWTLKHTVSHLHLFGTDKQHFGYDDKVVSIHPRNIIFL
VSLNDGIFISYEMDSREVHYILLVNHATRQWVVLPESARSDKRQIAYLGFDPAVSSHF
HVFELVEKNPANADGEADDGELDATINALEILSSETGVWSHKDIGWHQIGVLDDWRS
VFFNGMLHLITMGYVVAVVDVSGNSWRTIPMPGTLDDPDCNVDDGWFVDLSQGRLYFVN
TDRYDLYNSLSVWHLQDYSSDQWTLKHTVSHLHLFGRRRKDFGHDY"

/gene="OSJNBa0029B02.2"
/note="OSJNBa0029B02.2"
gene
    /gene="putative fibrillarin protein"
    join(10735. .10750,10831. .11466. .11561,11686. .11763,
    11800. .12024,13309. .13377,13413. .13667)
CDS
    /gene="OSJNBa0029B02.2"
    /note="similar to rice EST CR288939"
    /codon_start=1
    /product="putative fibrillarin protein"
    /protein_id="AAT77341.1"
    /db_xref="GI:50511418"
    /translation="MRPRGRFGRGRGRGGRGGGRGGRGGRGGRGGSGRGGRGGRGGSGV
    RGGRTPRGRGRGGRGGRGGRGGMKGGSKVVVPHHKHDGVFIAKAKEDALCTKNMVPGESV
    YGEKRISVQNEDGTKVERVRVWNPFRSKLAAAAVLGGVDNIWIAPGTRVLYLGGASGTTV
    SHVSDIVGPSPRYAILPFANQVILQTGLVIYAVEFSHRSGRDLVNMAKRTNVIPIIED
    ARHPARYRMLVGMWDVIFSDVAQPDQARILALNASYFLKNGGHFVISTKIIILIEVEI
    GEILDVILGIGLKLFEQANCIDSTMPAEAVFASEVEKLKADQFKPSEQVTLFEFPERDH
    ACVVGGYRMPKQKGTS"

CDS
    complement(17134. . .22580)
    /gene="OSJNBa0029B02.3"
    /note="OSJNBa0029B02.3"
gene
    /product="putative cellulose synthase"
    complement(join(17134. .17721,17941. .18291,18434. .18630,
    18791. .19049,19123. .19335,19413. .19538,19626. .19763,
    19843. .20108,20289. .20894. .21083,21167. .21251,
    21352. .21539,21729. .21924,22494. .22580))
    /gene="OSJNBa0029B02.3"
    /note=" similar to rice EST CF323300, AK102140, AK099228,
    AK098978, AK067967'"
    /codon_start=1
    /db_xref="GI:50511419"
    /protein_id="AAT77342.1"
    /translation="MAANAGMVAGSRNRNEFVWIRPDGDAPPAKPGKSVNGQVCQIC
    GDTVGVSARTGDVFVACNECAFPVCRPCYEYERKEGNQCCPQCKTRYKRHKGSPRVQGD
    EREDVDDLDNENTKHGNGKGPEWQIQRQGEDVDLSSSRHEQHRIPRLTGGCISG
    EIPDASPDRHSIRSGTSGSYVDPSVPVRIVDPSKDLNSYGINSVDWQERVASWRNKQ
    DKNMQVANKYPEARGGDMEGTGSNGEDMQMVDDARLPLSRIYPIPSNQLNLYRIVIVI
    LRILIMPFFQYRVTHYPRDAYGLMLVSVICHIFWPALSMLSLQAIESSETGVWPINRETYLDR
    LALRYDRKGEPSQLAPIDVFVSTVDPLKEPPLITANTVLSILAVDYPVDKVSCTVSDD
    GSAMLTFEALSETAEFARKWVPFCKKHNIEPRAPEFYFAQKIDYLKDKIQPSFVKERR
    AMKREYEEFKVRINALVAKAQKVPEEGWTMADGTAWPGMIQVFLGHSGGL
    DTDGNELPRLVVYSREKRPGFQHHKKAGAMNALIRVSAVLTNGAYLLNVDCDHYFNSS
    KALREAMCFMMDPALGRKTCYVQFPQRFDGIDLHDRYANRNIVFFDINMKGLDGIQGP
    VYGTGCCFNRQALYGIDGDVLTEADLEPNIVVAKSCCCGRCKKKSKSYMDSKRNMMKRTE
    SSAPIFNMEDIEBGYEDERSVLMSQKRLERFGQSPIFIASTFMTQGGIPPSTNP
    ASILKEAIHVISCCYEDKTEWGKEIGWIYGSVTEDILTGFKMARGWISIYCMPPRPC
    FKGSAPINLSDRLNQVLRWALGSVEILLSRHCPIWYGYNGRLKLLERLAYINIVYPI
    TSIPLIAYCVLPAICLLTNKFIIPEISNYAGMFFILFASIFATGILRLRWSGVGIED
    WWRNEQFWVIGGTSAHLFAVFQGLLKVLAGIDTNFVTVSKASDEDGDFAELYVFKWTS
    LLIPPTTVLIVINVGMVAGISYAINSGYGVWGPLFGKLPGKLFFSFIWILHYPFLKGLMGR
    QNRTPTIVIVWSILLASIFSLLWVKIDPFISFPTQKAVALGQCGVNC"

Qy   581
Db   84560
Qy   641
Db   84500
Qy   701
Db   84442
Qy   760
Db   84383

Db   84382 TGCAGTTGCTAGAAGCTCCCCAAATAGGCCCCTTAG 84347

```
RESULT 13
AP004323/c
LOCUS       AP004323    95209 bp    DNA    linear   HTG 21-MAR-2002
DEFINITION  Oryza sativa (japonica cultivar-group) chromosome 6 clone
            OJ1118_C02, *** SEQUENCING IN PROGRESS ***.
ACCESSION   AP004323
VERSION     AP004323.1  GI:16930108
KEYWORDS    HTG; HTGS PHASE2.
SOURCE      Oryza sativa (japonica cultivar-group)
ORGANISM    Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 95209)
  AUTHORS   Sasaki,T., Matsumoto,T. and Yamamoto,K.
  TITLE     Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, BAC
            clone:OJ1118_C02
  JOURNAL   Published Only in Database (2001)
REFERENCE   2  (bases 1 to 95209)
  AUTHORS   Sasaki,T., Matsumoto,T. and Yamamoto,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-NOV-2001) Takuji Sasaki, National Institute of
            Agrobiological Sciences, Rice Genome Research Program; Kannondai
            2-1-2, Tsukuba, Ibaraki 305-8602, Japan
            (E-mail:tsaaaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
            Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT     The nucleotide sequence of this BAC clone was generated by
            combining Monsanto and RGP-Japan sequencing data.
            NOTE: It currently consists of 1 contigs. Gaps between the contigs
            are represented as runs of N. The order of the pieces is believed
            to be correct as given, however the order of the pieces between them
            sequence will be replaced by the finished sequence as soon as it is
            available and the accession number will be preserved.
            NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source          1..95209
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="6"
                     /clone="OJ1118_C02"
ORIGIN

Query Match         8.5%;  Score 121.6;  DB 2;  Length 95209;
Best Local Similarity   74.0%;  Pred. No. 1.7e-17;
Matches  188;  Conservative  0;  Mismatches  49;  Indels  17;  Gaps  2;

Qy         592 GGGAGCTTAAAATATGAAAAAGCAGCTGCTGAGAAGCTACTGGGTGAGAATCTGAAGAAT 651
               ||||||| ||  |  |||  ||||||| || |||||||||  |||||||||| |||| ||
Db       89229 GGGAGCTTAAGATTTGAGAAGCAGCTAGCTGGTGAGAAGCTAGCTGGTGAGAATCTGGAGAAG 89170

Qy         652 TTGA----------GTTCTACGTTCATTCTCCAGATTCTTACAATTACAGATTCTT 696
               | ||          |||| |  |||| ||| ||| ||| | ||| ||||| ||||
Db       89169 CTGGAAAACCCAGCTTCTCTGGCTTCTCAGATTCTCACAACTACAGATTCTC 89110

Qy         697 ATAAATTTAGGTAAAAAAGCTGGACTGTTT--GGGGAGGCTTCTGTGTCAGCGGAGGAGATTCGTGA 754
               | AAA T   GGTAAAAAAGCTGGACTGTTT  |||GAGGCTTCTGA     CA CTGGGAGGATTC TAGTA
Db       89109 AGAAACTGGGGTAAAAAAGCTGGACTGTTTCGGGGAGGCTTCTGACTGTTCTGACACTGGGAGGATTCCTAGTA 89050

Qy         755 GAAGCTGCAGCTGCTAGAAGCTTCCCCAAACAGACCCCTAGTTGTACTCTCGAGCTGATCGA 814
               |||||||||||||||||||||| ||||||||||| ||||| |       |   |    ||
Db       89049 GAAGCTGCAGCTGCTAGAAGCTCCCCAAACAGGCCCCATARGACAACCTACTCTCCTCTT 88990

Qy         815 TTCACTCTATTTTA 828
               ||| || |   ||
Db         815                     
```

Db   88989 TGGGCTGCAGTTAA 88976

```
RESULT 14
AP005966/c
LOCUS       AP005966   187273 bp    DNA    linear   PLN 28-JUL-2004
DEFINITION  Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 6,
            BAC clone:B1047H05.
ACCESSION   AP005966
VERSION     AP005966.3  GI:50726539
KEYWORDS    .
SOURCE      Oryza sativa (japonica cultivar-group)
ORGANISM    Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   Sasaki,T., Matsumoto,T. and Katayose,Y.
  TITLE     Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, BAC
            clone:B1047H05
  JOURNAL   Published Only in Database (2002)
REFERENCE   2  (bases 1 to 187273)
  AUTHORS   Sasaki,T., Matsumoto,T. and Katayose,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-NOV-2002) Takuji Sasaki, National Institute of
            Agrobiological Sciences, Rice Genome Research Program; Kannondai
            2-1-2, Tsukuba, Ibaraki 305-8602, Japan
            (E-mail:tsaaaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
            Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT     On Jul 27, 2004 this sequence version replaced gi:47971628.
            Genes were predicted from the integrated results of the following:
            GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH
            (http://www.softberry.com/), GeneMark.hmm
            (http://opal.biology.gatech.edu/GeneMark/), GlimmerM
            (http://www.tigr.org/tdb/glimmerm/glmr_form.html), RiceHMM
            (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor
            (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4
            (http://globin.cse.psu.edu/html/docs/sim4.html), gap2
            (http://www.tigr.org/software/glimmerm/), BLASTN and BLASTX. The
            genomic sequence was searched against NCBI NonRedundant Protein
            database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
            sequence database at RGP or DDBJ. Protein homologies of the coding
            regions were searched against NCBI NonRedundant Protein database
            with BLASTP. ESTs represent the identified cDNA sequences using
            BLASTN with the corresponding DDBJ accession no. and RGP clone ID.
            Full-length cDNAs represent the identified cDNA sequences using
            BLASTN with the corresponding DDBJ accession no.
            A gene with identity or significant homology to a protein is
            classified based on the protein name to indicate the homology level
            such as same name, 'putative-', and '-like protein'. A gene without
            significant homology to any protein but with full-length cDNA or
            EST homology (covering almost the entire length of partial
            sequence) is classified as an 'unknown' protein. A gene predicted
            by two or more gene prediction programs is classified as a
            'hypothetical' protein according to IRGSP standard. A gene
            predicted by a single gene prediction program is also classified as
            a probable 'hypothetical' protein and is included as a
            miscellaneous feature of the sequence.
            The orientation of the sequence is from M13rev to -21M13 of the BAC
            clone. This sequence of B1047H05 clone has an overlap with P0633B08
            clone. Detailed information on overlap and
            (DDBJ: AP003622) at 3' end. Detailed information of this entry is
            assembly quality together with annotation of this entry is
            available at http://rgp.dna.affrc.go.jp/GenomeSeq.html.
FEATURES             Location/Qualifiers
     source          1..187273
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="6"
                     /clone="B1047H05"
     gene            complement(327..2093)
                     /gene="B1047H05.1"
```

```
misc_feature   complement(327..2093)
               /gene="B1047H05.1"
               /note="probably inactive due to 5' exon missing in CDS"

gene           join(2836..2984,3392..3462,3591..3666,3754..3827,
               4065..4180)
               /gene="B1047H05.2"
               /pseudo
               /note="pseudogene, GAG-POL precursor"

mRNA           join(2836..2984,3392..3462,3591..3666,3754..3827,
               4065..4180)
               /gene="B1047H05.2"

misc_feature   join(<7570..>8024
               /gene="B1047H05.3"
               /note="supported by full-length cDNA(s): AK062314"

gene           complement(8076..10056,10155..11416))
               /gene="B1047H05.4"

mRNA           complement(8076..10056,10155..11416))
               /gene="B1047H05.4"

misc_feature   complement(11918..13366)
               /gene="B1047H05.5"
               /note="retrotransposon protein-like"

gene           complement(join(<11918..11979,12095..12218,13015..13123,
               13260..>13366))
               /gene="B1047H05.5"

mRNA           complement(join(<11918..11979,12095..12218,13015..13123,
               13260..>13366))
               /gene="B1047H05.5"
               /note="start and end point are not identified"

CDS            complement(join(11918..11979,12095..12218,13015..13123,
               13260..13366))
               /gene="B1047H05.5"
               /note="similar to Oryza sativa chromosome 1, P0463F06.28"
               /codon_start=1
               /product="hypothetical protein"
               /protein_id="BAD34174.1"
               /db_xref="GI:50726540"
               /translation="MPDGYPLEDGVGRLACMRRRSTACAPPVCHHPHSLLLLGLRREV
               KQLKVTELVAPHRETAQALLCQFGRGLRCKPHKGSGHLGPRGRGRSNPTPPRGNQEEG
               GHLTLGPDVPEGRKRSGHMTPFIKPGDRWDRSP"

gene           complement(15812..17991)
               /gene="B1047H05.6"

mRNA           complement(join(<15812..15981,17970..>17991))
               /gene="B1047H05.6"
               /note="start and end point are not identified"

CDS            complement(join(15812..15981,17970..17991))
               /gene="B1047H05.6"
               /note="predicted by GENSCAN etc."
               /codon_start=1
               /product="hypothetical protein"
               /protein_id="BAD34175.1"
               /db_xref="GI:50726541"
               /translation="MEKNSERGWTRFTAWAALRPNPCEQLVFAHGGGGERYQEHVQRV
               GVVIGBGEHVVVQWTRTLM"

gene           18451..18813
               /gene="B1047H05.7"

mRNA           join(18451..18560,18678..>18813)
               /gene="B1047H05.7"
               /note="start and end point are not identified"

CDS            join(18451..18560,18678..18813)
               /gene="B1047H05.7"
               /note="predicted by GeneMark.hmm etc."
               /codon_start=1
               /product="hypothetical protein"
               /protein_id="BAD34176.1"
               /db_xref="GI:50726542"
               /translation="MEAASVVTAFLGEMNVVRPVRGVAFNELAXRFGMHNLPAKMAQ
```

```
               QGLKTNDNEVKKGWGPTATHIDGVBGQRLCGPAARTG"

gene           20823..22640
               /gene="B1047H05.8"

mRNA           <20823..>22640
               /gene="B1047H05.8"
               /note="start and end point are not identified"

CDS            20823..22640
               /gene="B1047H05.8"
               /codon_start=1
               /product="putative disease resistance protein Cf-2.1"
               /protein_id="BAD34177.1"
               /db_xref="GI:50726543"
               /translation="MSSLGFSSPGFLQLLYILKFFCLLPLVIGSNETSSDRQALLCLK
               SQLTGSAAVLSSWSNASMRFCSWHGVTCSTQYPRRVTALDLSESEGITGSISPCIANLT
               VLTKLQLSNNSFYGSIPSELGFLTQLSIILNISMNSLBGNIPSELTSCFKLQKIDLSNN
               KLQGSIPSAFGDLTELRTLILTSNRLSGDIPQSLGSNLSLTVDLGRNALAGRIPQSL
               ASSTSLQFLILTSNTLSGRLPKALLNSSLIFLDLCQNNFVGSIPPVTAISPKMYYLD
               LRFNHLTGTIPSSLGNLSSLTYLCLIGNNLVGSIPDTLGHVPTLETLAVNVNNLSGPV
               PSIFNVLSLTYLGMAWNSLTGRLPSNIGYTLPNIQQLLLPNNKFSGGSIPSSLLNASH
               LQRLFLTNNSFTGHIPFGSLQNLEILDMAYNMLBAGDWSFVSSLTNCSKLTQLLLDG
               NNLQGNLPSCIGNLSSSLEHLMLRNNMISGLIPPGIGNLKSLANTLYMDDNYLTGNIPP
               TIGYLIHNMNKLYMDYNLTGNIPPTIGYLFSFSHNRLSGGQIPGTIGNLVQLNE
               LRLDENNLSGSIPASIRHCTQLIKLNLAHNSLHGTTIAS"

gene           24538..24975
               /gene="B1047H05.9"

mRNA           <24538..>24975
               /gene="B1047H05.9"
               /note="start and end point are not identified"

CDS            24538..24975
               /gene="B1047H05.9"
               /codon_start=1
               /note="predicted by GlimmerM etc."
               /product="hypothetical protein"
               /protein_id="BAD34178.1"
               /db_xref="GI:50726544"
               /translation="MDLEEGRGLERRGGGEGPALSSVAASSSLLTEPPSLLVELDGCR
               RLIACADGRRLLFRRCSSRPLLABSDGCCRLLACANGRRLLIFRRHRHRSSWPLLABS
               DGCRQLLVCSNGHRLLLFHSRHRICRCRLLARAKPPTATGSSSR"

gene           complement(26909..27955)
               /gene="B1047H05.10"

mRNA           complement(join(<26909..27047,27707..27833,27868..>27955))
               /gene="B1047H05.10"
               /note="start and end point are not identified"

CDS            complement(join(26909..27047,27707..27833,27868..27955))
               /gene="B1047H05.10"
               /note="predicted by RiceHMM etc."
               /codon_start=1
               /product="hypothetical protein"
               /protein_id="BAD34179.1"
               /db_xref="GI:50726545"
               /translation="MMRLFTTAPSARVAVLRAAVAVVPMYGGSRAADPRKVAHLDSTV
               HVNNPTPRHGHKTQGADKAAEAADPKSESKPPLGSSGTKCKTSVHYKVVWTGAYMFSA
               SPMVSTVLRPARKII"

gene           complement(30838..31158)
               /gene="B1047H05.11"

mRNA           complement(<30838..>31158)
               /gene="B1047H05.11"
               /note="start and end point are not identified"

CDS            complement(30838..31158)
               /gene="B1047H05.11"
               /note="start and end point are not identified"
```

```
Query Match         8.5%;    Score 121.6;   DB 8;    Length 187273;
Best Local Similarity  74.0%;   Pred. No. 1.8e-17;
Matches  188;  Conservative   0;  Mismatches  49;  Indels  17;  Gaps  2;

Qy    592  GGGGAGCTTAAAAATTAATGAAAAGCAGCTGCTGAGAAGCTAGCTGGTGGTGAGAATCTGAAGAAT  651
           |||||||| |||  ||| ||| ||||| ||||||||||||||| |||||||||||||   |||
Db  122593  GGGGAGCTTAAGATTTTGGAAGAAGCAGCTGCTGAGAAGCTAGCTGGTGGTGAGAAGTCTGGAGAAG  122534

Qy    652  TTGA------------------GTTCTAGTTCATTCTCCAGATTCTACAAATTACAGATTCTT  696
                                 ||  |  ||||||  |||  ||||| ||||||||| ||||
Db  122533  CTGGAAAAACCCAGCTTCTGGCTTCTAGTTCATTTTTCAGATTCTACAACTACAGATTCTC  122474

Qy    697  ATAATTTAGGTAAAAAGCTGGACTGTGTT--GGGAGCTTCGTCAGCCGGAGATTCGTGA  754
           |  |  |||||    ||||| |||| |     |||||||||  |  ||| ||| |
Db  122533                                                            122533
```

```
Db   122473  AGAATCTGGGTAAAAAGCTGGACTGTTTGGGGGGAGCTTCTGACTGAGATTCTAGTA 122414
             ||||||||||||  ||||||||||||||  |||   | | ||||||||||||||||||
Qy      755  GAAGCTGCAGCTGCCTAGAAGCTTCCCCAAAACAGACACCCTAGTTGTACTTCTAGCTGATCGA 814
             |||||||||||||  ||||||||||||||
Db   122413  GAAGCTGCAGCTGCTAGAAGTCTCCCCAAAACAGGCCCCATAATGACAACCTACTACTCTCCCTCTT 122354

Qy      815  TTCACTCTATTTTA 828
             || ||| ||||
Db   122353  TGGGCTGCAGTTAA 122340
```

```
RESULT 15
LOCUS      AP003625/c
DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 6, PAC
           *** SEQUENCING IN PROGRESS ***,
ACCESSION  AP003625
VERSION    AP003625.1  GI:14020963
KEYWORDS   HTG; HTGS_PHASE2.
SOURCE     Oryza sativa (japonica cultivar-group)
ORGANISM   Oryza sativa (japonica cultivar-group)
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
           Ehrhartoideae; Oryzeae; Oryza.
REFERENCE  1
AUTHORS    Sasaki,T. and Yamamoto,K.
TITLE      Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC
           clone:P0701F09
JOURNAL    Published Only in Database (2001)
REFERENCE  2  (bases 1 to 134159)
AUTHORS    Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE      Direct Submission
JOURNAL    Submitted (10-MAY-2001) Takuji Sasaki, National Institute of
           Agrobiological Resources, Rice Genome Research Program; Kannondai
           2-1-2, Tsukuba, Ibaraki 305-8602, Japan
           (E-mail:tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
           Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT    NOTE: It currently consists of 1 contigs. Gaps between the contigs
           are represented as runs of N. The order of the pieces is believed
           to be correct as given, however the sizes of the gaps between them
           are based on estimates that have provided by the submitter. This
           sequence will be replaced by the finished sequence as soon as it is
           available and the accession number will be preserved.
           * NOTE: This is a 'working draft' sequence.
           * This sequence will be replaced
           * by the finished sequence as soon as it is available and
           * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source          1..134159
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="6"
                     /clone="P0701F09"
ORIGIN
```

LOCUS   AP003625   134159 bp   DNA   linear   HTG 21-MAR-2002
                                   chromosome 6 clone P0701F09,

```
Query Match          8.4%;  Score 120;  DB 2;  Length 134159;
Best Local Similarity  73.6%;  Pred. No. 4.2e-17;
Matches 187;  Conservative  0;  Mismatches 50;  Indels 17;  Gaps  2;

Qy      592  GGGAGCTTAAAATTATGAAAAAGCAGCTAGCTAGCTGGTGGTAGAGATCTGAAGAAT 651
             |||||||||||   ||||||||||||||||||||||  |||  | ||||||||||||
Db    57995  GGGAGCTTAAGATTTTGAGAGAAGCAGCTAGCTGCTGAGAGAGGCTAGCTGGTGTGAGAATCTGAGAAG 57936

Qy      652  TTGA----------------GTTCTACGTTCATTCTCCCAGATTCTACAATTACAGATTCTT 696
             |||                 |||  |   ||  || || ||  || ||   || ||||||||
Db    57935  CTGGAAAACCCAGCTTCTCGGCTTCTAGTTCATTTTTCAGATTCTACAACTACACAGATTCTC 57876

Qy      697  ATAATTTAGGTAAAAAGCTGGACTGTTT---GGGAGCTTCTGTGTCAGGCCGAGATTCTGTGA 754
             ||||||  |||||||||||||||||||||   |||||||||||   |||||  || ||||||||
Db    57875  AGAATCTGGGCTAAAAAGCTGGACTGTTTTGGGGGGAGCTTCTGACAACTGGAGATTCTAGTA 57816
```

```
Qy      755  GAAGCTGCAGCTGCCTAGAAGCTTCCCCAAAACAGACACCCTAGTTGTACTTCTAGCTGATCGA 814
             |||||||||||||  |||||||||||||
Db    57815  GAAGCTGCAGCTGCTAGAAGCTCCCCCAAAACAGGCCCCATAATGACAACCTACTACTCTCCCTCTT 57756

Qy      815  TTCACTCTATTTTA 828
             || ||| ||||
Db    57755  TGGGCTGCAGTTAA 57742
```

Search completed: September 23, 2005, 18:23:22
Job time : 4320 secs

This Page Blank (uspto)

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: September 23, 2005, 12:19:35 ; Search time 577 Seconds
           (without alignments)
           14732.666 Million cell updates/sec

Title: US-10-602-166-2
Perfect score: 1436
Sequence: 1 cattcagaatcatctccagc...........tatcaggtggtcaacctgg 1436
Scoring table: IDENTITY_NUC
           Gapop 10.0 , Gapext 1.0
Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :
        1: N_Geneseq_16Dec04:*
        1: geneseqn1980s:*
        2: geneseqn1990s:*
        3: geneseqn2000s:*
        4: geneseqn2001as:*
        5: geneseqn2001bs:*
        6: geneseqn2002as:*
        7: geneseqn2002bs:*
        8: geneseqn2003as:*
        9: geneseqn2003bs:*
       10: geneseqn2003cs:*
       11: geneseqn2003ds:*
       12: geneseqn2004as:*
       13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	Match DB	ID	Description
1	1436	100.0	1436	3	AAA27334	Aaa27334 Rice RA8
2	1436	100.0	303	3	AAA27333	Aaa27333 Rice RA8
3	1196	83.3	1196	3	AAA27335	Aaa27335 Rice RA8
4	240	16.7	240	3	AAA27341	Aaa27341 Rice RA8
5	95	6.6	1008	3	AAA27338	Aaa27338 Rice RA8
6	90.4	6.3	2000	3	ADA71441	Ada71441 Rice gene
7	90.4	6.3	2000	12	ADJ41616	Adj41616 Plant cDN
8	77.2	5.4	17953	6	AAD38802	Aad38802 Rice RGA8
c 9	71	4.9	2000	8	ADA72788	Ada72788 Rice gene
c 10	69.2	4.8	2000	8	ADA72600	Ada72600 Rice gene
c 11	60.4	4.2	2000	8	ADA71536	Ada71536 Rice gene
c 12	59.2	4.1	2000	8	ADA73357	Ada73357 Rice gene
13	58.8	4.1	2000	12	ADJ40892	Adj40892 Plant cDN
14	57.6	4.0	11460	6	AAD38805	Aad38805 CODR1 ORF
15	57.6	4.0	69300	6	AAD38804	Aad38804 BAC clone
16	57.6	4.0	91552	6	AAD38803	Aad38803 BAC clone
c 17	55.8	3.9	2000	12	ADJ40653	Adj40653 Plant cDN
c 18	53	3.7	2000	8	ADA71743	Ada71743 Rice gene
c 19	50.8	3.5	6239	8	ABK31185	Abk31185 Signal tr
c 20	50.8	3.5	6239	6	ABL70146	Abl70146 Chemical1
21	50.8	3.5	6239	6	AAS61072	Aas61072 Human gen
22	50.6	3.5	11534	6	ABL32343	Abl32343 Human imm
23	50.4	3.5	6735	6	ABL33374	Abl33374 Human imm
24	50.4	3.5	8056	8	ABZ10246	Abz10246 Haematopo
25	49.8	3.5	914	13	ADR65796	Adr65796 Cotton cD
26	49.8	3.5	2889	12	ADN94933	Adn94933 Indian ri
27	49.8	3.5	12763	6	ABL32303	Abl32303 Human imm
28	49.6	3.5	2000	8	ADA72032	Ada72032 Rice gene
29	49.6	3.5	5518	8	ABZ10115	Abz10115 Haematopo
30	49.6	3.5	5518	10	ADE84213	Ade84213 Human lym
31	49.6	3.5	17131	6	ABL33052	Abl33052 Human imm
c 32	49.2	3.4	2000	8	ADA71938	Ada71938 Rice gene
33	49	3.4	11049	6	ABL32668	Abl32668 Human imm
34	49	3.4	11049	6	ABL92218	Abl92218 Chemical1
35	49	3.4	11049	6	ABL49321	Abl49321 Human pol
c 36	48.8	3.4	8056	8	ABZ10100	Abz10100 Haematopo
37	48.4	3.4	7104	6	ABL34589	Abl34589 Human met
38	48.4	3.4	7104	6	ABL70396	Abl70396 Chemical1
39	48.4	3.4	7104	7	ADS99850	Ads99850 Complemen
c 40	48.4	3.4	16766	6	ABL34157	Abl34157 Human imm
c 41	48.4	3.4	1537	8	ADA72364	Ada72364 Rice gene
42	48	3.3	5518	6	ABL32216	Abl32216 Human imm
43	48	3.3	5518	6	ABL54317	Abl54317 Chemical1
44	48	3.3	5518	6	ABN80018	Abn80018 Human che
45	48	3.3	5518	8	ABZ09969	Abz09969 Haematopo

ALIGNMENTS

RESULT 1
AAA27334
ID   AAA27334 standard; DNA; 1436 BP.
XX
AC   AAA27334;
XX
DT   10-AUG-2000 (first entry)
XX
DE   Rice RA8 anther-specific gene regulator.
XX
KW   RA8; rice; anther-specific; self-pollination; antisense; transgene;
KW   transgenic plant; wheat; maize; orchardgrass; Sorghum bicolor;
KW   male-sterile plant; ds.
XX
OS   Oryza sativa.
XX
FH   Key   Location/Qualifiers
FH
FT   exon   1..1288
FT   /*tag= e
FT   exon   1..1196
FT   /number= 1
FT   promoter   1116..1119
FT   /*tag= a
FT   CAAT_signal   1116..1119
FT   /*tag= b
FT   TATA_signal   1145..1151
FT   /*tag= c
FT   CDS   1247..2769
FT   /*tag= d
FT   /product= "RA8"
FT   /note= "Contains 2 introns"
FT   /partial
FT   intron   1289..1422
FT   /*tag= f
FT   /number= 1
FT   exon   1423..1436
FT   /*tag= g
FT   /number= 2
XX
PN   WO200026389-A2.
XX
PD   11-MAY-2000.
XX
PF   02-NOV-1999; 99WO-EP008360.

```
XX
PR   03-NOV-1998;   98KR-00046973.
PR   19-NOV-1998;   98KR-00050126.
XX
PA   (NOVS )  NOVARTIS AG.
PA   (NOVS )  NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX
PI   An G, Jeon J, Chung Y, Lee S;
XX
DR   WPI; 2000-365632/31.
DR   P-PSDB; AAY96273.
XX
PT   Novel promoters of anther-specific transcription used to create
PT   transgenic male-sterile plants.
XX
PS   Claim 15; Page 28-29; 32pp; English.
XX
CC   The present sequence is the gene regulator for the rice RA8 protein. This
CC   protein is expressed in an anther-specific manner in the plant, where it
CC   aids in the development of the structure. The gene can be used to create
CC   transgenic plants which do not possess a properly formed anther, and thus
CC   are male-sterile. This is useful as it prevents self-pollination, which
CC   will then aid breeding and hybrid seed production. In addition to rice,
CC   this process can also be used in wheat, maize, orchardgrass and Sorghum
CC   bicolor
XX
SQ   Sequence 1436 BP; 387 A; 313 C; 273 G; 463 T; 0 U; 0 Other;

Query Match             100.0%;  Score 1436;  DB 3;  Length 1436;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1436; Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;
```

(sequence alignment block — query/database sequences in full)

```
RESULT 2
AAA27333
ID   AAA27333 standard; DNA; 3003 BP.
XX
AC   AAA27333;
XX
DT   10-AUG-2000  (first entry)
XX
DE   Rice RA8 anther-specific gene.
XX
KW   RA8; rice; anther-specific; self-pollination; antisense; transgene;
KW   transgenic plant; wheat; maize; orchardgrass; Sorghum bicolor;
KW   male-sterile plant; ds.
XX
OS   Oryza sativa.
XX
FH   Key                 Location/Qualifiers
```

```
FT   exon            1..1288
FT                   /*tag= e
FT                   /number= 1
FT   promoter        1..1196
FT                   /*tag= a
FT   CAAT_signal     1116..1119
FT                   /*tag= b
FT   TATA_signal     1145..1151
FT                   /*tag= c
FT   CDS             1247..2769
FT                   /*tag= d
FT                   /product= "RA8"
FT                   /note= "Contains 2 introns"
FT   intron          1289..1422
FT                   /*tag= f
FT                   /number= 1
FT   exon            1423..1555
FT                   /*tag= g
FT                   /number= 2
FT   intron          1556..2149
FT                   /*tag= h
FT                   /number= 2
FT   exon            2150..3003
FT                   /*tag= i
FT                   /number= 3
XX
PN   WO20026389-A2.
XX
PD   11-MAY-2000.
XX
PF   02-NOV-1999;  99WO-EP008360.
XX
PR   03-NOV-1998;  98KR-0004697.3.
PR   19-NOV-1998;  98KR-0005012.6.
XX
PA   (NOVS ) NOVARTIS AG.
PA   (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX
PI   An G, Jeon J, Chung Y, Lee S;
XX
DR   WPI; 2000-365632/31.
DR   P-PSDB; AAY96273.
XX
PT   Novel promoters of anther-specific transcription used to create
PT   transgenic male-sterile plants.
XX
PS   Claim 3; Page 27-28; 32pp; English.
XX
CC   The present sequence is the gene for the rice RA8 protein. This protein
CC   is expressed in an anther-specific manner in the plant, where it aids in
CC   the development of the structure. The gene can be used to create
CC   transgenic plants which do not possess a properly formed anther, and thus
CC   are male-sterile. This is useful as it prevents self-pollination, which
CC   will then aid breeding and hybrid seed production. In addition to rice,
CC   this process can also be used in wheat, maize, orchardgrass and Sorghum
CC   bicolor
XX
SQ   Sequence 3003 BP; 736 A; 747 C; 694 G; 826 T; 0 U; 0 Other;

Query Match     100.0%;  Score 1436;  DB 3;  Length 3003;
Best Local Similarity   100.0%;
Matches 1436;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy        1 CATTCAGAATCATCTCCAGCCTACAAATGTACTCGTCTCCCATAAATACAAGTGTCTCTATGA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 CATTCAGAATCATCTCCAGCCTACAAATGTACTCGTCTCCCATAAATACAAGTGTCTCTATGA 60

Qy       61 TTCAAAATTTGTCCTACAAATATAAACATTTCCAGCATGAAATCCATACATTAATTTTCAG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 TTCAAAATTTGTCCTACAAATATAAACATTTCCAGCATGAAATCCATACATTAATTTTCAG 120

Qy      121 CTAATCAGATGCTTGGAGGGAAAAAACTTAAGCGGATTCAATATGCAAAAAATTGATCACTGA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 CTAATCAGATGCTTGGAGGGAAAAAACTTAAGCGGATTCAATATGCAAAAAATTGATCACTGA 180
```

```
Qy    1261  CGCCATCGGCCATCGCCCATGGCTCTCCATGGTGTGCAGAGATATCCAGCATGGAAATTAACT  1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261  CGCCATCGCCATCGCCCATGGCTCTCCATGGTGTGCAGAGATATCCAGCATGGAAATTAACT  1320

Qy    1321  TGACGCCATATCATCTCATCATCACTTGCATTTCAACTTCTGGATTGTGTGCAGACATGTTTA  1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321  TGACGCCATATCATCTCATCATCACTTGCATTTCAACTTCTGGATTGTGTGCAGACATGTTTA  1380

Qy    1381  GCTGATCAGTAAAGTTGCGTGTGTCCTTATCAGGTGGTTGTTCAAACCTGG  1436
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381  GCTGATCAGTAAACGTTGCCGTGTTGAATTGGTCCTTATCAGGTGGTTCAAACCTG  1436
```

RESULT 3

ID   AAA27335 standard; DNA; 1196 BP.
XX
AC   AAA27335;
XX
DT   10-AUG-2000 (first entry)
XX
DE   Rice RA8 anther-specific gene promoter.
XX
KW   RA8; rice; anther-specific; self-pollination; antisense; transgene;
KW   transgenic plant; wheat; maize; orchardgrass; Sorghum bicolor;
KW   male-sterile plant; ds.
XX
OS   Oryza sativa.
XX
PN   WO200026389-A2.
XX
PD   11-MAY-2000.
XX
PF   02-NOV-1999; 99WO-EP008360.
PR   03-NOV-1998; 98KR-0046973.
PR   19-NOV-1998; 98KR-00050126.
XX
PS   Claim 9; Page 29; 32pp; English.
XX
CC   The present sequence is the gene promoter for the rice RA8 protein. This
CC   protein is expressed in an anther-specific manner in the plant, where it
CC   aids in the development of the structure. The gene can be used to create
CC   transgenic plants which do not possess a properly formed anther, and thus
CC   are male-sterile. This is useful as it prevents self-pollination, which
CC   will then aid breeding and hybrid seed production. In addition to rice,
CC   this process can also be used in wheat, maize, orchardgrass and Sorghum
CC   bicolor.
XX
PT   Novel promoters of anther-specific transcription used to create
PT   transgenic male-sterile plants.
XX
PA   (NOVS) NOVARTIS AG.
PA   (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX
PI   An G, Jeon J, Chung Y, Lee S;
XX
DR   WPI; 2000-365632/31.
XX
SQ   Sequence 1196 BP; 336 A; 242 C; 224 G; 394 T; 0 U; 0 Other;

Query Match   83.3%; Score 1196; DB 3; Length 1196;
Best Local Similarity   100.0%; Pred. No. 8.9e-290;
Matches 1196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1  CATTCAGAATCATCTCCAGCCTACAAATGTACTTCTCTCTATGATAATACAAGTGTCTCTATGA  60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1  CATTCAGAATCATCTCCAGCCTACAAATGTACTTCTCTCTATGATAATACAAGTGTCTCTATGA  60

Qy      61  TTCAAAATTTGTCCTTACAATATAAAACATTTCCAGCAATGAAAATCCATACACATTAATTTCAG  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61  TTCAAAATTTGTCCTTACAATATAAAACATTTCCAGCAATGAAAATCCATACACATTAATTTCAG  120

Qy     121  CTAATCAGATGGCTTGGAGGGAGGGAAAAATCTAAGGCGGATTCAATATGCAAAAAAATTGATCACTGA  180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121  CTAATCAGATGGCTTGGAGGGAGGGAAAAATCTAAGGCGGATTCAATATGCAAAAAAATTGATCACTGA  180

Qy     181  AGTAACTGAAAGAGAATATCTCGTTTTAACATTAGTGCTAGTATTTATTAAACAACTAAA  240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181  AGTAACTGAAAGAGAATATCTCGTTTTAACATTAGTGCTAGTATTTATTAAACAACTAAA  240

Qy     241  AAATTGTTTATATTTTTTAGTTACAAAACGAGTAGTACCAGTAGCAGAGAGGATCG  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241  AAATTGTTTATATTTTTAGTTACAAAATCGAGTAGTACCAGTAGCAGAGAGGATCG  300

Qy     301  TGTTCCGGATCACCTGAGAGAAACCGTCAGGTGGTTGTTCTGTTGGTTGGAGCTTA  360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301  TGTTCCGGATCACCTGAGAAACCGTCAGGTCGGTTTCTGTGTGGCCCAGCCGTCCAGCGGATCGAGAAT  360

Qy     361  TCGGAGGATCCGCCGTCGTTTCTTCTTCCTGAAAATCTGCAAGTCCCAGCGACGAGCAG  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361  TCGGAGGATCCGCCGTCGTTTCTTCTTCCTGAAAATCTGCAAGTCCCAGCGACGAGCAG  420

Qy     421  AGCAAGAGCAAATGGGGCGTGCAGGGGAGTTTGATACTTTGATCCACTAGCTACTAGCTAGGCG  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421  AGCAAGAGCAAATGGGGCGTGCAGGGGAGTTTGATACTTTGATCCACTAGCTACTAGCTAGGCG  480

Qy     481  TTCGTTCCAGTCGGCTTCTACAATTACAGATTCTTATAATTAGGTAAAAAGCTGGACT  540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481  TTCGTTCCAGTCGGCTTCTACAATTACAGATTCTTATAATTAGGTAAAAAGCTGGACT  540

Qy     541  AAGATTATATTCCTCACATTTTTCTTCCTATGCTGCTAGTCGTCGTTGTTCGTTGGGAGCTTA  600
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541  AAGATTATATTCCTCACATTTTTCTTCCTATGCTGCTAGTCGTCGTTGTTCGTTGGGAGCTTA  600

Qy     601  AAATTATGAAAAGCAGCTGCTGAGAGCTAGCTGCTGAGAATCTGAAGAATTTGAGTTCT  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601  AAATTATGAAAAGCAGCTGCTGAGAGCTAGCTGCTGAGAATCTGAAGAATTTGAGTTCT  660

Qy     661  ACGGTTCATTCTCCAGATTCTACAATTACAGATTCTTATAATTAGGTAAAAAGCTGGACT  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661  ACGGTTCATTCTCCAGATTCTACAATTACAGATTCTTATAATTAGGTAAAAAGCTGGACT  720

Qy     721  GTTTGGGAGCTTCTGTCGTCGTGCAGCGGGAGATTCTGTGAGAGCTGCAGCTGCTAGGAAGCTTCCC  780
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721  GTTTGGGAGCTTCTGTCGTCGTGCAGCGGGAGATTCTGTGAGAGCTGCAGCTGCTAGGAAGCTTCCC  780

Qy     781  CAAACAGACCCCTAGTTGTACTCGATCGATCGACTCACTCTTATTTTATTATACACCTTGC  840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781  CAAACAGACCCCTAGTTGTACTCGATCGATCGACTCACTCTTATTTTATTATACACCTTGC  840

Qy     841  TCTCTAGCTTATCAAACGTAGCCAAGACTTGAAATTTTGAAAGCTAAAATGATTTGATGT  900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841  TCTCTAGCTTATCAAACGTAGCCAAGACTTGAAATTTTGAAAGCTAAAATGATTTGATGT  900

Qy     901  TCTTTTCATCGTAATTCACTTACCGACCTTAGTCGGCCATTGAAATTTTGAAATTAATT  960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901  TCTTTTCATCGTAATTCACTTACCGACCTTAGTCGGCCATTGAAATTTTGAAATTAATT  960

Qy     961  TTAGAGCTGATTTGAATTTATTTTCAGCGGGAAATTTATTTTCACGTATGTAAAGTTTTA  1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961  TTAGAGCTGATTTGAATTTATTTTCAGCGGGAAATTTATTTTCACGTATGTAAAGTTTTA  1020

Qy    1021  CCTATAAAATTAATTTAATTTCAGCGGGTAAGCATTAGTGTTTATGGGTTATGGTTTATATATCTGG  1080
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021  CCTATAAAATTAATTTAATTTCAGCGGGTAAGCATTAGTGTTTATGGGTTATGGTTTATATATCTGG  1080

Qy    1081  TATGCTTAAATCTCTTTACTTGGACTTAGTTGGGACAATTGTAAGCATTCTCGTTGCCA  1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081  TATGCTTAAATCTCTTTACTTGGACTTAGTTGGGACAATTGTAAGCATTCTCGTTGCCA  1140

Qy    1141  TCTCTTATAATACGGGCCTGCTAGCCTTTGCTCTTGTATCTGACACAAGAACTAGCTG  1196
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141  TCTCTTATAATACGGGCCTGCTAGCCTTTGCTCTTGTATCTGACACAAGAACTAGCTG  1196
```

RESULT 4

```
AAA27341
ID    AAA27341 standard; DNA; 240 BP.
XX
AC    AAA27341;
XX
DT    10-AUG-2000  (first entry)
XX
DE    Rice RA8 anther-specific gene fragment.
XX
KW    RA8; rice; anther-specific; self-pollination; antisense; transgene;
KW    transgenic plant; wheat; maize; orchardgrass; Sorghum bicolor;
KW    male-sterile plant; ds.
XX
OS    Oryza sativa.
XX
FH    Key             Location/Qualifiers
FH
FT    exon            1..92
FT                    /*tag= b
FT                    /number= 1
FT    CDS             51..240
FT                    /*tag= a
FT                    /product= "RA8"
FT                    /note= "Contains an intron"
FT    intron          93..226
FT                    /*tag= c
FT                    /number= 1
FT    exon            227..240
FT                    /*tag= d
FT                    /number= 2
XX
PN    WO20026389-A2.
XX
PD    11-MAY-2000.
XX
PF    02-NOV-1999;  99WO-EP008360.
XX
PR    03-NOV-1998;  98KR-0046973.
PR    19-NOV-1998;  98KR-0050126.
XX
PA    (NOVS ) NOVARTIS AG.
PA    (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX
PI    An G, Jeon J, Chung Y, Lee S;
XX
DR    WPI; 2000-365632/31.
DR    P-PSDB; AAY96273.
XX
PT    Novel promoters of anther-specific transcription used to create
PT    transgenic male-sterile plants.
XX
PS    Claim 13; Page 32; 32pp; English.
XX
CC    The present sequence is part of the gene for the rice RA8 protein. This
CC    protein is expressed in an anther-specific manner in the plant, where it
CC    aids in the development of the structure. The gene can be used to create
CC    transgenic plants which do not possess a properly formed anther, and thus
CC    are male-sterile. This is useful as it prevents self-pollination, which
CC    will then aid breeding and hybrid seed production. In addition to rice,
CC    this process can also be used in wheat, maize, orchardgrass and Sorghum
CC    bicolor. This sequence was used to create an expression vector for the
CC    protein.
XX
SQ    Sequence 240 BP; 51 A; 71 C; 49 G; 69 T; 0 U; 0 Other;
```

```
Qy          1197 GCAAAGTCTCTCAAGGGGGAACCGCCTCCATCTTCTCTTCCAGCTCCTCCCATGGCGTCCC 1256
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             1 GCAAAGTCTCTCAAGGGGGAACCGCCTCCATCTTCTCTTCCAGCTCCTCCCATGGCGTCCC 60
```

Query Match   16.7%; Score 240; DB 3; Length 240;
Best Local Similarity   100.0%;
Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1257 TCGTCGGCCATCGCCATCGGCCATGGGCCTCATGGTGCAGAGGATATCCAGCATGCTGAAATT 1316
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            61 TCGTCGGCCATCGCCATCGGCCATGGGCCTCATGGTGCAGAGGATATCCAGCATGCTGAAATT 120

Qy          1317 AACTTGACGCAATATCATCTCATCATCACTTGCATTTCAACTTTCTGGATTGTGCAGACATG 1376
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           121 AACTTGACGCAATATCATCTCATCATCACTTGCATTTCAACTTTCTGGATTGTGCAGACATG 180

Qy          1377 TTTAGCTGATCAGTAAACGTTGCCGTGTTGAATTGGTGTCCTTATCAGGTGGTTCAACCTGG 1436
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           181 TTTAGCTGATCAGTAAACGTTGCCGTGTTGAATTGGTGTCCTTATCAGGTGGTTCAACCTGG 240
```

RESULT 5

```
AAA27338
ID    AAA27338 standard; cDNA; 1008 BP.
XX
AC    AAA27338;
XX
DT    10-AUG-2000  (first entry)
XX
DE    Rice RA8 anther-specific gene.
XX
KW    RA8; rice; anther-specific; self-pollination; antisense; transgene;
KW    transgenic plant; wheat; maize; orchardgrass; Sorghum bicolor;
KW    male-sterile plant; ss.
XX
OS    Oryza sativa.
XX
FH    Key             Location/Qualifiers
FH
FT    CDS             51..845
FT                    /*tag= a
FT                    /product= "RA8"
XX
PN    WO20026389-A2.
XX
PD    11-MAY-2000.
XX
PF    02-NOV-1999;  99WO-EP008360.
XX
PR    03-NOV-1998;  98KR-0046973.
PR    19-NOV-1998;  98KR-0050126.
XX
PA    (NOVS ) NOVARTIS AG.
PA    (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX
PI    An G, Jeon J, Chung Y, Lee S;
XX
DR    WPI; 2000-365632/31.
DR    P-PSDB; AAY96273.
XX
PT    Novel promoters of anther-specific transcription used to create
PT    transgenic male-sterile plants.
XX
PS    Claim 19; Page 31; 32pp; English.
XX
CC    The present sequence is the gene for the rice RA8 protein. This protein
CC    is expressed in an anther-specific manner in the plant, where it aids in
CC    the development of the structure. The gene can be used to create
CC    transgenic plants which do not possess a properly formed anther, and thus
CC    are male-sterile. This is useful as it prevents self-pollination, which
CC    will then aid breeding and hybrid seed production. In addition to rice,
CC    this process can also be used in wheat, maize, orchardgrass and Sorghum
CC    bicolor.
XX
SQ    Sequence 1008 BP; 145 A; 360 C; 325 G; 178 T; 0 U; 0 Other;
```

Query Match   6.6%; Score 95; DB 3; Length 1008;
Best Local Similarity   100.0%;
Matches   95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1197 GCAAAGTCTCTCAAGGGGGAACCGCCTCCATCTTCTCTTCCAGCTCCTCCCATGGCGTCCC 1256
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             1 GCAAAGTCTCTCAAGGGGGAACCGCCTCCATCTTCTCTTCCAGCTCCTCCCATGGCGTCCC 60
```

Db
Qy   1 GCAAAGTCCTCAAGGCGGAACGGCCGTTCCATCTTCTCCTTCCAGCTCCTCCCATGGGCGTCCC 60
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1257 TCGTCGCCATCGCCATCGGCTCTCAATGGGTG 1291
        ||||||||||||||||||||||||||||||||
Db   61 TCGTCGCCATCGCCATCGGCTCTCAATGGGTG 95

RESULT 6
ID   ADA71441 standard; DNA; 2000 BP.
XX
AC   ADA71441;
XX
DT   20-NOV-2003 (first entry)
XX
DE   Rice gene, SEQ ID 4764.
XX
KW   Plant; bacterial infection; fungal infection; viral infection; rice;
KW   gene; ds.
XX
OS   Oryza sativa.
XX
PN   WO200300898-A1.
XX
PD   03-JAN-2003.
XX
PF   22-JUN-2001; 2001WO-IB001105.
XX
PR   22-JUN-2001; 2001WO-IB001105.
XX
PA   (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI   Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI   Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
DR   WPI; 2003-175290/17.
XX
PT   Identifying at least one gene involved in plant resistance or response to
PT   pathogenic infection for conferring resistance or tolerance to a plant to
PT   bacterial, fungal or viral infection by determining or detecting plant
PT   gene expression.
XX
CC   Claim 27; SEQ ID NO 4764; 899pp; English.
XX
CC   The present invention relates to a method (M1) for identifying genes
CC   involved in plant resistance or response to pathogenic infection. M1
CC   comprises identifying a gene whose expression is significantly altered in
CC   the incompatible interaction of plant gene expression relative to
CC   expression of the gene in an uninfected plant, in a mutant plant that
CC   does not express a gene associated with response to pathogenic infection,
CC   or in a corresponding incompatible or compatible interaction. (M1) is
CC   useful for conferring resistance to resistance or tolerance to a plant to
CC   bacterial, fungal or viral infection. The present sequence was used to
CC   illustrate the invention.
XX
SQ   Sequence 2000 BP; 630 A; 343 C; 386 G; 640 T; 0 U; 1 Other;

Query Match   6.3%; Score 90.4; DB 8; Length 2000;
Best Local Similarity   71.1%; Pred. No. 3.8e-12;
Matches 155; Conservative   0; Mismatches   46; Indels   17; Gaps   2;

Qy 593 GGAGCTTAAATATATGAAAAAGCAGCTGCCTGAGAAGCTAGCTGGGTGAGAAGCTGAAGAATT 652
        ||||||||| || ||||||||||| || |||||||||||||| ||||||||| ||||
Db 1686 GGAGCTTAAGATTCTGAGAAGCAACTGCTAAGAAGCTAGCTGGTGAGAAGCTGGAGAAGC 1745

Qy 653 TGAGTTCTAC--------------------GTTCATTCTCCAGATTCTACAATTACAGATTCTTA 697
        || |||||||| |||||||||||||| || ||||||||
Db 1746 TGGAAAACCAGCTTCTGTCTTCTACAGTTCATTTTCCAGATTCTACAATCTATAGATTCTTA 1805

Qy 698 TAATTTAGGTAAAAAAGCTGACTGGTTTGG--GAGCTTCTGTCAGCCGGAGATTCGTCGTGAG 755
        ||| ||||||||||||| ||| ||| || |||||||| || ||| |||| |||| ||
Db 1806 GAATCTTGGTAAAAAAGATGGGTTGTTAGGGAGGAGCTTTGGTAGCTAGAGAATTATAGGAG 1865

Qy 756 AAGCTGCAGCTGCTAGAAGCTTCCCAAACAGACCCCT 793
        |||| |||||| || |||||||| |||| ||||
Db 1866 AAGATGCAACTGCTAAAAACTCCTTAAATAGGTCCTT 1903

RESULT 7
ID   ADJ41616 standard; cDNA; 2000 BP.
XX
AC   ADJ41616;
XX
DT   06-MAY-2004 (first entry)
XX
DE   Plant cDNA #2616.
XX
KW   Plant; gene; ss; transcription; plant genome augmentation; cereal;
KW   soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;
KW   maize; barley; sorghum; wheat; crop plant; insecticide resistance;
KW   stress tolerance; salt tolerance; cold tolerance; drought tolerance;
KW   plant nutrition; apical dominance; dwarfism; early flowering; antiviral;
KW   antifungal.
XX
OS   Eukaryota.
XX
PN   US2004016025-A1.
XX
PD   22-JAN-2004.
XX
PF   26-SEP-2002; 2002US-00260238.
XX
PR   26-SEP-2001; 2001US-0325277P.
PR   26-SEP-2001; 2001US-0325448P.
PR   04-APR-2002; 2002US-0370620P.
XX
PA   (BUDW/) BUDWORTH P.
PA   (MOUG/) MOUGHAMER T.
PA   (BRIG/) BRIGGS S P.
PA   (COOP/) COOPER B.
PA   (GLAZ/) GLAZEBROOK J.
PA   (GOFF/) GOFF S A.
PA   (KATA/) KATAGIRI F.
PA   (KREP/) KREPS J.
PA   (PROV/) PROVART N.
PA   (RICK/) RICKE D.
PA   (ZHUT/) ZHU T.
XX
DR   WPI; 2004-190374/18.
XX
PT   New rice promoter, useful for manipulating crop plants to alter or
PT   improve phenotypic characteristics, e.g. produce large quantities of oil
PT   or proteins, resistance to insecticides, virus or fungi, stress tolerance
PT   or high nutritional value.
XX
PS   Claim 1; SEQ ID NO 2616; 230pp; English.
XX
PI   Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;
PI   Goff SA, Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;
XX
CC   The invention relates to plant nucleotide sequences that direct seed-,
CC   leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential
CC   or constitutive transcription of an operatively linked nucleic acid
CC   segment. The invention also relates to a method for augmenting a plant
CC   genome and a method of identifying a gene, where its expression is
CC   altered in the seed, leaf, stem, panicle, pollen, root or is constitutive
CC   in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,
CC   canola, cotton, peanut, tobacco or sugar beet, preferably maize,
CC   sorghum, rice or wheat. The polynucleotides and the polypeptides they
CC   encode are useful for manipulating crop plants to alter or improve
CC   phenotypic characteristics, to produce large quantities of oil or
CC   proteins, to incur resistance to insecticides, viruses or fungi, and to
CC   incur stress tolerance (e.g. salt, cold or drought) to ensure the plants
CC   have a high nutritional value with reduced apical dominance or dwarfism,

CC   early flowering or altered metabolic pathways. This sequence represents a
CC   plant nucleic acid of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification but was obtained in
CC   electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ   Sequence 2000 BP; 628 A; 343 C; 387 G; 641 T; 0 U; 1 Other;

Query Match   6.3%; Score 90.4; DB 12; Length 2000;
Best Local Similarity 71.1%; Pred. No. 3.8e-12;
Matches 155; Conservative   0; Mismatches 46; Indels 17; Gaps 2;

Qy   593 GGAGCTTAAAATTATGAAAAGCAGCGTGCTGAGAAGCTAGCGTGGTGAGAAGTCGAGAATT 652
          |||||||||| ||| ||||| ||||||||||||||||||||| |||||||||||| ||||
Db   1688 GGAGCTTAAGAATTCTGAGAAGCAAACTGCTAAGAAGCTAGCTGGTGAGAAGTCGGAGAAGC 1747

Qy   653 TGGAGTTCTAC------------GTTCAATTCTCCCAGATTCTACAAATTACAGAGATTCTTA 697
          |||| ||| |||||| ||||||| |||||||||||||||||||||
Db   1748 TGGAAAACCCAGCTTCTGTGTCTTCAGTTCATTTTCATTTCTACAATTCTACATTCTTA 1807

Qy   698 TAATTTAGGTAAAAAAGCTGGACTGTTTGG--GAGCTTCTGTCAGCCGGAGATTCGTGAG 755
          |||||||||||||||| |||||||||||||||||||| |||||
Db   1808 GAATCTTGGTAAAAAGATGGGTTGTTGTTAGGGAGGAGCTTTTGGTAGCTAGAGATTATAGGAG 1867

Qy   756 AAGCTGCAGCGTGCTAGAAGCTTCCCCAAACAGACCCCT 793
          |||| ||||||||||||||||||||||||||||
Db   1868 AAGATGCAAACTGCTAAAAAACTCCTTTAAATAGGTCCTT 1905

RESULT 8
AAD38802
ID   AAD38802 standard; DNA; 17953 BP.
XX
AC   AAD38802;
XX
DT   23-SEP-2002 (first entry)
XX
DE   Rice RGA8 contig.30Nippon DNA.
XX
KW   Rice; chromosome 11; Indica rice cultivar; CO39; avirulence gene;
KW   AVR1-CO39; rice blast pathogen; transgenic plant; plant breeding;
KW   resistance; agricultural; horticultural; plant protectant; ds.
XX
OS   Oryza sativa.
XX
PN   WO20023 49927-A2.
XX
PD   02-MAY-2002.
XX
PF   19-OCT-2001; 2001WO-US046331.
XX
PR   20-OCT-2000; 2000US-0242313P.
PR   09-JUL-2001; 2001US-0303897P.
XX
PA   (WISC) WISCONSIN ALUMNI RES FOUND.
PA   (USDA) US DEPT OF AGRICULTURE.
XX
PI   Leong SA, Farman ML, Chauhan RS, Durfee TJ;
XX
DR   WPI; 2002-471442/50.
XX
PT   New PiCO39(t) polynucleotides from chromosome 11 of Indica rice cultivar
PT   CO39 useful for conferring or improving resistance of plants to strains
PT   of Magnaporthe grisea or other plant pathogens with avirulence gene AVR1-
PT   CO39.
XX
PS   Example 2; Page 59-68; 175pp; English.
XX
CC   The invention relates to a polynucleotide isolated from chromosome 11 of
CC   Indica rice cultivar CO39, flanked by marker R2316 and RG1094 comprising
CC   one or more genes that confer resistance to strains of Magnaporthe grisea
CC   useful for conferring or improving resistance of plants to strains of the
CC   rice blast pathogen, Magnaporthe grisea and other pathogens. The Pi-

CC   CO39(t) nucleic acids may be used as probes to detect the presence of
CC   and/or expression of Pi-CO39(t) genes, and to produce large quantities of
CC   pure Pi-CO39(t) proteins. Purified gene products of Pi-CO39(t) are useful
CC   in producing polyclonal or monoclonal antibodies useful as sensitive
CC   detection reagents for the presence and accumulation of Pi-CO39(t)
CC   polypeptides. The transgenic plants are useful for plant breeding or
CC   directly in agricultural or horticultural applications. The present
CC   sequence is rice RGA8 contig.30Nippon DNA. This sequence comprises
CC   portion of BAC clone 82N20 from Nipponbare, containing RGA8 sequence and
CC   ORF NBR7
XX
SQ   Sequence 17953 BP; 5074 A; 3615 C; 3742 G; 5522 T; 0 U; 0 Other;

Query Match   5.4%; Score 77.2; DB 6; Length 17953;
Best Local Similarity 68.3%; Pred. No. 1.8e-08;
Matches 142; Conservative   0; Mismatches 53; Indels 13; Gaps 2;

Qy   589 TTTGGGAGCTTAAAAATTATGAAAAGCAGCGTGCTGAGAAGCTAGCGTGGTGAGAAGTCGAA 648
          |||||||||| |||| ||| ||||||||||||||||||||| |||||| |||| ||||| ||
Db 17088 TTGGGGAGCTTAAGAATTCTGAGAAGTAGCTGGTGAGTGAGCTAGCCTTCTGATAAATCTGGAA 17147

Qy   649 AATTTGAGTTCTACGTTCATTCTCCAGATTCTACAAATTACAGAGATTCTTATAAATTTAGGTA 708
          |||| ||||| || |||||||| |||||||| |||| ||||||||||
Db 17148 AAGTTCGGGTTT----TTCAACTTCGTCGGCTTCTATTAACTACAGATTCTTAGAACCTGAGTG 17203

Qy   709 AAAAGCTGGACTGTTTGGGGAGCTTCGTGTCAGCCGGAGAGTTCGTGTGAGAAGCTGCAGCTGC 768
          ||||| ||||| ||||| || |||| |||| ||||| ||||||| ||
Db 17204 AGAATGTAGACTATTTGAGG---------AGCTGGAAATTCTAAAAGAAAACTGCAAAATAC 17254

Qy   769 TAGAAGCTTCCCAAACAGACCCCTTAGT 796
          ||||||||| || |||||| |||||||
Db 17255 TAGAAGCTTCACTCAAAACAGGGCCCTTAGT 17282

RESULT 9
ADA72788/c
ID   ADA72788 standard; DNA; 2000 BP.
XX
AC   ADA72788;
XX
DT   20-NOV-2003 (first entry)
XX
DE   Rice gene, SEQ ID 6113.
XX
KW   Plant; bacterial infection; fungal infection; viral infection; rice;
KW   gene; ds.
XX
OS   Oryza sativa.
XX
PN   WO200300898-A1.
XX
PD   03-JAN-2003.
XX
PF   22-JUN-2001; 2001WO-IB001105.
XX
PR   22-JUN-2001; 2001WO-IB001105.
XX
PA   (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI   Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI   Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
DR   WPI; 2003-175290/17.
XX
PT   Identifying at least one gene involved in plant resistance or response to
PT   pathogenic infection for conferring resistance or tolerance to a plant to
PT   bacterial, fungal or viral infection by determining or detecting plant
PT   gene expression.
XX
PS   Claim 27; SEQ ID NO 6113; 899pp; English.
XX
CC   The present invention relates to a method (M1) for identifying genes
CC   involved in plant resistance or response to pathogenic infection. M1

CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction,
CC bacterial, fungal or viral infection. The present sequence was used to
CC useful for conferring resistance or tolerance to resistance to a plant to
CC illustrate the invention.
XX
SQ Sequence 2000 BP; 497 A; 452 C; 457 G; 594 T; 0 U; 0 Other;

Query Match 4.9%; Score 71; DB 8; Length 2000;
Best Local Similarity 66.4%; Pred. No. 2.8e-07;
Matches 144; Conservative 0; Mismatches 50; Indels 23; Gaps 2;

```
Qy 589 TTTGGGAGCTTAAAATTATGAAAAGCAGCTGCTGAGAAGCTAGCTGGTGAGAAATCTGAAG 648
        || ||||| | |||||||| ||||||| | ||  | ||| |||  ||||| ||||| ||| 
Db 859 TTGGGGAGGTGTAAGATTATGAGAAGCAGAGANGGTNGGTAGCCAGCTTCTGAGAATCGGAA 800

Qy 649 AATTTGAGTTCTACG------------TTCATTCTCCAGATTCTACAANTTACAGATT 693
        |||| ||| |            |||| || || | |||| || ||||  | ||| |||
Db 799 AAGTTAGATTTCCCAACTTATNGGCTTCTATTTTATTTCTAGAATTCTACAACTACAGTTT 740

Qy 694 CTTATAATTTAGGTAAAAAAGCTAGACAGCTGGACTGTTTGGGAGCTTCTGCTCAGCCGGA 753
        ||| ||| | ||| |||||||||| |  ||| || ||||| ||||          ||  ||
Db 739 CTTAGAATCTAGACCAAAAAGCTGACTGTTTAGGAGGAGATTTCT------AATTCTAGG 688

Qy 754 AGAAGCTGCAGCCTGCTAGGAAGCTTCCCAAACAGACC 790
        |||||||||  | ||| ||  |||| ||| ||| | ||
Db 687 AGAAGCTGCTGCAGTTAGAAAACTCCCCAACAAGCC 651
```

RESULT 10
ADA72600/c
ID ADA72600 standard; DNA; 2000 BP.
XX
AC ADA72600;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 5925.
XX
KW Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
XX
OS Oryza sativa.
XX
PN WO200300898-A1.
XX
PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX
PR 22-JUN-2001; 2001WO-IB001105.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
DR WPI; 2003-175290/17.
XX
PT gene expression.
PT bacterial, fungal or viral infection by determining or detecting plant
PT involved in plant resistance or response to pathogenic infection. M1
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT Identifying at least one gene involved in plant resistance or response to
XX
PS Claim 27; SEQ ID NO 5925; 899pp; English.
XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to

CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance or tolerance to resistance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 2000 BP; 617 A; 367 C; 428 G; 586 T; 0 U; 2 Other;

Query Match 4.8%; Score 69.2; DB 8; Length 2000;
Best Local Similarity 76.6%; Pred. No. 7.9e-07;
Matches 98; Conservative 0; Mismatches 28; Indels 2; Gaps 1;

```
Qy 664 TTCATTCTCCAGATTCTACAANTTACAGATTCTTATAATTTAGGTAAAAAGCTGGACTGTT 723
        ||  || | | ||||||||||  ||| |||||| | |||| |||| |||| | ||| ||||
Db 985 TTTATTTTCTGAATTCTACAACTACGGATTCTCAGAATCTCGGGTGAAAAATATAAACTGTT 926

Qy 724 T--GGGAGCTTCTGCTCAGCCGGAGATTCTGTGAGAAGCTGCAGCTGCTAGAAGCTTCCC 781
        |  |||||  ||| || |||| |||| ||||   |||||| ||| | |  || || |||||
Db 925 TGAGGGAGTTTCTACCAGCTGGAGAATTCTACGGAGAAGCTGCAGTTGCCAGAAGCTCCCCT 866

Qy 782 AAACAGAC 789
        ||| || |
Db 865 AAATAGGC 858
```

RESULT 11
ADA71536
ID ADA71536 standard; DNA; 2000 BP.
XX
AC ADA71536;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 4859.
XX
KW Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
XX
OS Oryza sativa.
XX
PN WO200300898-A1.
XX
PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX
PR 22-JUN-2001; 2001WO-IB001105.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
DR WPI; 2003-175290/17.
XX
PT gene expression.
PT bacterial, fungal or viral infection by determining or detecting plant
PT involved in plant resistance or response to pathogenic infection. M1
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT Identifying at least one gene involved in plant resistance or response to
XX
PS Claim 27; SEQ ID NO 4859; 899pp; English.
XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction,
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC bacterial, fungal or viral infection. The present sequence was used to
CC useful for conferring resistance or tolerance to resistance to a plant to
CC illustrate the invention.

```
XX
SQ   Sequence 2000 BP; 653 A; 376 C; 363 G; 607 T; 0 U; 1 Other;

XX
Qy   Query Match     4.2%; Score 60.4; DB 8; Length 2000;
     Best Local Similarity   62.5%; Pred. No. 0.0013;
     Matches 115; Conservative   0; Mismatches 61; Indels   8; Gaps   1;

Qy   608 GAAAAGCAGCTGCTGAGGAAGCTAGCTCGGTGAGGAGAATTGAGTTCTTACGTTCA 667
         ||||| |||||||||||||||||||||||||||||||||||||||||
Db   200 GTAACCCAGTTCTCGAGAAATCTGGATAAAGCTGGGTTTTCCAGTTTCGGCTTCAGTTCA 259

Qy   668 TTCTCCAGATTCTACAATTACAGATTCTTATATAATTTAGGTAAAAAAGCTGGACTGTTTGGG 727
         ||||| |||||||||||||     ||||||||||||||||||||||||||
Db   260 TTTTCTCGGATTCTACAACTATAACTTCTTAGAATTCTGAACCAAAAGCTAAATTGTTTGGG 319

Qy   728 AGCTTCTTGTCACCCGAGAGATTCTGTGAGAAGCTGCAGCTGCTCTAGAGAGCTTCCCAAACAG 787
         ||||||||        |||||||||||||||||||||||||||||
Db   320 ATAGTTTCT-------GATTATGGGAGAGAAGCTGCAGCGAGCTAAAAACTCTTCCAAACAG 371

Qy   788 ACCC 791
         ||||
Db   372 GCCC 375

RESULT 12
ID   ADA73357/c
AC   ADA73357
XX
DT   20-NOV-2003 (first entry)
XX
DE   Rice gene, SEQ ID 6683.
XX
KW   Plant; bacterial infection; fungal infection; viral infection; rice;
KW   gene; ds.
XX
OS   Oryza sativa.
XX
PN   WO200300898-A1.
XX
PD   03-JAN-2003.
XX
PF   22-JUN-2001; 2001WO-IB001105.
XX
PR   22-JUN-2001; 2001WO-IB001105.
XX
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI   Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI   Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
DR   WPI; 2003-175290/17.
XX
PT   gene expression.
PT   bacterial, fungal or viral infection by determining or detecting plant
PT   pathogenic infection for conferring resistance or tolerance to a plant to
PT   Identifying at least one gene involved in plant resistance or response to
XX
PS   Claim 27, SEQ ID NO 6683; 899pp; English.
XX
CC   The present invention relates to a method (M1) for identifying genes
CC   involved in plant resistance or response to pathogenic infection. M1
CC   comprises identifying a gene whose expression is significantly altered in
CC   the incompatible interaction of plant gene expression relative to
CC   expression of the gene in an uninfected plant, in a mutant plant that
CC   does not express a gene associated with response to pathogenic infection,
CC   or in a corresponding incompatible or compatible interaction,
CC   useful for conferring resistance or tolerance to resistance to a plant to
CC   bacterial, fungal or viral infection. The present sequence was used to
CC   illustrate the invention.
XX
SQ   Sequence 2000 BP; 618 A; 438 C; 283 G; 660 T; 0 U; 1 Other;


Qy   Query Match     4.1%; Score 59.2; DB 8; Length 2000;
     Best Local Similarity   63.8%; Pred. No. 0.00026;
     Matches 134; Conservative   0; Mismatches 53; Indels  23; Gaps  2;

Qy   590 TTGGGAGCTTAAAATTATAGAAAAGCAGCTGCTTGAGAAGCTAGCTGGGTGAGAATCTGAAGA 649
         ||||| |||||| | ||||||||| |||||| |||||| ||| |||||||| ||| |||||
Db  1592 TTGGGTGCTTAAGAGTTCTAAAAAATCTGCTGCTGTCGGTATCCCAGCTTTATGAGAAATGGAA 1533

Qy   650 ATTTGAGTTCTAC------------GTTCATTCTCCAGATTCTGGGGTTGTAGTTTATTTTTTTGGATTTATAACTACAGATC 694
         |                       |||||||| |||||||| |||||||||||||||||||||||||||||||||||||||
Db  1532 AGCTGGGGTTTTCTTGCTTCTGGGGTTGTAGTTTATTTTTTTGGATTTATAACTACAGATC 1473

Qy   695 TTATAAATTTAGGTAAAAAAGCTGGACTGTTTGGGGAGCGTTCTGTCGTCCCGAGAGATTCTGTGA 754
         | ||||||| ||| |||||||||||||||| |||         ||||||||||| ||||| ||
Db  1472 TCAGAATCTGGGTGATNAATCTGGACTGTTTGAGAG------AACTGGAGATTCTGGTA 1421

Qy   755 GAAGCTGCAGCTGCTAGAAGCTTCCCCAA 784
         ||| |||| || |||||||||||||| ||
Db  1420 AAATCTGCAACAACTAGAAAAATCCCCTCAA 1391

RESULT 13
ID   ADJ40892
AC   ADJ40892
XX
DT   06-MAY-2004 (first entry)
XX
DE   Plant cDNA #1892.
XX
KW   Plant; gene; ss; transcription; plant genome augmentation; cereal;
KW   soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;
KW   maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;
KW   stress tolerance; salt tolerance; cold tolerance; drought tolerance;
KW   plant nutrition; apical dominance; dwarfism; early flowering; antiviral;
KW   antifungal.
XX
OS   Eukaryota.
XX
PN   US2004016025-A1.
XX
PD   22-JAN-2004.
XX
PF   26-SEP-2002; 2002US-00260238.
XX
PR   04-APR-2002; 2002US-0370620P.
PR   26-SEP-2001; 2001US-0325448P.
PR   26-SEP-2001; 2001US-0325277P.
XX
PA   (BUDW/) BUDWORTH P.
PA   (MOUG/) MOUGHAMER T.
PA   (BRIG/) BRIGGS S P.
PA   (COOP/) COOPER B.
PA   (GLAZ/) GLAZEBROOK J.
PA   (GOFF/) GOFF S A.
PA   (KATA/) KATAGIRI F.
PA   (KREP/) KREPS J.
PA   (PROV/) PROVART N.
PA   (RICK/) RICKE D.
PA   (ZHUT/) ZHU T.
XX
PI   Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;
PI   Goff SA, Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;
XX
DR   WPI; 2004-190374/18.
XX
PT   New rice promoter, useful for manipulating crop plants to alter or
PT   improve phenotypic characteristics, e.g. produce large quantities of oil
PT   or proteins, resistance to insecticides, virus or fungi, stress tolerance
PT   or high nutritional value.
XX
```

PS   Claim 33; SEQ ID NO 1892; 230pp; English.
XX
CC   The invention relates to plant nucleotide sequences that direct seed-,
CC   leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential
CC   or constitutive transcription of an operatively linked nucleic acid
CC   segment. The invention also relates to a method for augmenting a plant
CC   genome and a method of identifying a gene, where its expression is
CC   altered in the seed, leaf, stem, panicle, pollen, root or is constitutive
CC   in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,
CC   canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,
CC   sorghum, rice or wheat. The polynucleotides and the polypeptides they
CC   encode are useful for manipulating crop plants to alter or improve
CC   phenotypic characteristics, to produce large quantities of oil or
CC   proteins, to incur resistance to insecticides, viruses or fungi, and to
CC   incur stress tolerance (e.g. salt, cold or drought) to ensure the plants
CC   have a high nutritional value with reduced apical dominance or dwarfism,
CC   early flowering or altered metabolic pathways. This sequence represents a
CC   plant nucleic acid of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification but was obtained in
CC   electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ   Sequence 2000 BP; 539 A; 473 C; 428 G; 559 T; 0 U; 1 Other;

Query Match   4.1%; Score 58.8; DB 12; Length 2000;
Best Local Similarity 66.5%; Pred. No. 0.00032;
Matches 141; Conservative 0; Mismatches 47; Indels 24; Gaps 3;

Qy 589 TTTGGGAGCTTAAAAATTATGAAAAACAGCTGCTGAGAAGCTAGCTGGTGGGTGAGAAGCTGAAG 648
Db 627 TTGGAGAGCTTAAAAATTCTGATAAGCAACTGATTAGTAGCCAACT--TTGCAAATCTGGAA 685

Qy 649 AATTTGAGTTCT----------ACGTTCATTCTCCAGATTCTTACAATTACAGATT 693
Db 686 AAGCTGGGTTTTCTTAGTTTTTGACTTCATTTTCTAGATTCTTACAAATACAGTTT 745

Qy 694 CTTATAATTTAGGTAAAAAGCTGGACTGTTTGGGAGCTTCGTCAGCCGGAGAATTCTGTG 753
Db 746 CTCAGAATCTGGACCAAAAAGCTAGACTGTTTTAGATAGTTTTT------GATTCTGAT 797

Qy 754 AGAAGCTGCAGCTGCTGTAGAAGCTTCCCAAAC 785
Db 798 AGAAGTTGCAACAGCCAAAAAGCTCTCCCAAAC 829

RESULT 14
AAD38805/c
ID   AAD38805 standard; DNA; 11460 BP.
XX
AC   AAD38805;
XX
DT   23-SEP-2002 (first entry)
XX
DE   CODR1 ORF from rice variety CO39.
XX
KW   Rice; chromosome 11; Indica rice cultivar; CO39; avirulence gene;
KW   AVR1-CO39; rice blast pathogen; transgenic plant; plant breeding;
KW   resistance; agricultural; horticultural; plant protectant; ds.
XX
OS   Oryza sativa.
XX
PN   WO200234927-A2.
XX
PD   02-MAY-2002.
XX
PF   19-OCT-2001; 2001WO-US046331.
XX
PR   20-OCT-2000; 2000US-0242313P.
PR   09-JUL-2001; 2001US-0303897P.
XX
PA   (WISC) WISCONSIN ALUMNI RES FOUND.
PA   (USDA) US DEPT OF AGRICULTURE.
XX
PI   Leong SA, Farman ML, Chauhan RS, Durfee TU;

XX
DR   WPI; 2002-471442/50.
XX
PT   New PiCO39(t) polynucleotides from chromosome 11 of Indica rice cultivar
PT   CO39 useful for conferring or improving resistance of plants to strains
PT   of Magnaporthe grisea or other plant pathogens with avirulence gene AVR1-
PT   CO39.
XX
PS   Claim 6; Page 143-148; 175pp; English.
XX
CC   The invention relates to a polynucleotide isolated from chromosome 11 of
CC   Indica rice cultivar CO39, flanked by marker R2316 and RG1094 comprising
CC   one or more genes that confer resistance to strains of Magnaporthe grisea
CC   having avirulence gene AVR1-CO39. The Pi-CO39(t) polynucleotides are
CC   useful for conferring or improving resistance of plants to strains of the
CC   rice blast pathogen, Magnaporthe grisea and other pathogens. The Pi-
CC   CO39(t) nucleic acids may be used as probes to detect the presence of
CC   and/or expression of Pi-CO39(t) genes, and to produce large quantities of
CC   pure Pi-CO39(t) proteins. Purified gene products of Pi-CO39(t) are useful
CC   in producing polyclonal or monoclonal antibodies useful as sensitive
CC   detection reagents for the presence and accumulation of Pi-CO39(t)
CC   polypeptides. The transgenic plants are useful for plant breeding or
CC   directly in agricultural or horticultural applications. The present
CC   sequence is CODR1 ORF from rice variety CO39
XX
SQ   Sequence 11460 BP; 3238 A; 2381 C; 2138 G; 3703 T; 0 U; 0 Other;

Query Match   4.0%; Score 57.6; DB 6; Length 11460;
Best Local Similarity 63.0%; Pred. No. 0.0013;
Matches 109; Conservative 0; Mismatches 56; Indels 8; Gaps 1;

Qy   638 AGAATCTGAAGAAATTGAGTTCTACGTTCATTCTCCAGATTCTTACAATTACAGATTCTTA 697
Db 4898 AGCTTCTCCAGCTTCTGGCCTTCTTAGTTCATTTTCTAGATTCTGTAACTACAGATTCTCA 4839

Qy   698 TAATTTAGGTAAAAAGCTGGACTGTTTGGGAGCTTCGTCAGCCGGAGAATTCTGTGAGAA 757
Db 4838 GAAGCTGTGGACTG------TTTGGGGTAGCTTCTAGCAGAAGAGCAGCTTTTGGGAAAA 4787

Qy   758 GCTGCAGCTGCTGTAGAAGCTTCCCAAAACAGACCCCTAGTTGTTGTACTTCTAGTGA 810
Db 4786 GTTACAGCTGGGACAAGCAAAGCTCCCCAAAACAGGACCTAAGTTACACAGTTGGTCA 4734

RESULT 15
AAD38804
ID   AAD38804 standard; DNA; 69300 BP.
XX
AC   AAD38804;
XX
DT   23-SEP-2002 (first entry)
XX
DE   BAC clone E2P5 from rice variety CO39.
XX
KW   Rice; chromosome 11; Indica rice cultivar; CO39; avirulence gene;
KW   AVR1-CO39; rice blast pathogen; transgenic plant; plant breeding;
KW   resistance; agricultural; horticultural; plant protectant; ds.
XX
OS   Oryza sativa.
XX
PN   WO200234927-A2.
XX
PD   02-MAY-2002.
XX
PF   19-OCT-2001; 2001WO-US046331.
XX
PR   20-OCT-2000; 2000US-0242313P.
PR   09-JUL-2001; 2001US-0303897P.
XX
PA   (WISC) WISCONSIN ALUMNI RES FOUND.
PA   (USDA) US DEPT OF AGRICULTURE.
XX
PI   Leong SA, Farman ML, Chauhan RS, Durfee TU;

```
XX
DR   WPI; 2002-471442/50.
XX
PT   New PiCO39(t) polynucleotides from chromosome 11 of Indica rice cultivar
PT   CO39 useful for conferring or improving resistance of plants to strains
PT   of Magnaporthe grisea or other plant pathogens with avirulence gene AVRl-
PT   CO39.
XX
PS   Example 4; Page 110-142; 175pp; English.
XX
CC   The invention relates to a polynucleotide isolated from chromosome 11 of
CC   Indica rice cultivar CO39, flanked by marker R2316 and RG1094 comprising
CC   one or more genes that confer resistance to strains of Magnaporthe grisea
CC   having avirulence gene AVR1-CO39. The Pi-CO39(t) polynucleotides are
CC   useful for conferring or improving resistance of plants to strains of the
CC   rice blast pathogen, Magnaporthe grisea and other pathogens. The Pi-
CC   CO39(t) nucleic acids may be used as probes to detect the presence of
CC   and/or expression of Pi-CO39(t) genes, and to produce large quantities of
CC   pure Pi-CO39(t) proteins. Purified gene products of Pi-CO39(t) are useful
CC   in producing polyclonal or monoclonal antibodies useful as sensitive
CC   detection reagents for the presence and accumulation of Pi-CO39(t)
CC   polypeptides. The transgenic plants are useful for plant breeding or
CC   directly in agricultural or horticultural applications. The present
CC   sequence is BAC clone E2P5 from rice variety CO39
XX
SQ   Sequence 69300 BP; 19720 A; 15094 C; 15006 G; 19380 T; 0 U; 100 Other;

Query Match             4.0%;  Score 57.6;  DB 6;  Length 69300;
Best Local Similarity   63.0%;  Pred. No. 0.0026;
Matches 109;  Conservative   0;  Mismatches 56;  Indels   8;  Gaps   1;

Qy        638 AGAATCTGAAGAAATTGAGTTCTACGTTCAATTCTCCAGATTCTACAATTACAGAATTCTTA 697
              |||||| ||| |||||||||||||||  ||||||| ||||| |||| || |||||||||||||| |
Db      12015 AGCTTCTCCAGCTTCTCGGCTTCTTAGTTCATTTTCAGATTCTGTAACTACAGAATTCTCA 12074

Qy        698 TAATTTAGGTAAAAAGCTGGACTGTTTGGGAGCTTCGTGTCAGCCGGAGATTCTGTGAGAA 757
              ||| |||||    |   |||||     |||| ||| || ||||| |||| || |||| ||||
Db      12075 GAAGCTGTGGACTG-----------TTTGGGGTAGCTTCTTAGCAGAAAGCAGCTTTTGGGAAAA 12126

Qy        758 GCTGCAGAGCTGCTTAGAAGCTTCCCCAAACAGACCCCTAGTTGTTGTACTCTTAGCTGA 810
              |||||||||||||||| ||||||  |||| ||||| | ||||| ||||||| |||||||| |
Db      12127 GTTACAGCTGGGACAAGCCTCCCCAAACACAGGACCTAAGTTACACACAGTTGGGTCA 12179

Search completed: September 23, 2005, 17:11:19
Job time : 583 secs
```

This Page Blank (uspto)